

NONLINEAR YOUNG INTEGRALS AND DIFFERENTIAL SYSTEMS IN HÖLDER MEDIA

YAOZHONG HU AND KHOA N. LÊ

ABSTRACT. For Hölder continuous random field $W(t, x)$ and stochastic process φ_t , we define nonlinear integral $\int_a^b W(dt, \varphi_t)$ in various senses, including pathwise and Itô-Skorohod. We study their properties and relations. The stochastic flow in a time dependent rough vector field associated with $\dot{\varphi}_t = (\partial_t W)(t, \varphi_t)$ is also studied and its applications to the transport equation $\partial_t u(t, x) - \partial_t W(t, x) \nabla u(t, x) = 0$ in rough media is given. The Feynman-Kac solution to the stochastic partial differential equation with random coefficients $\partial_t u(t, x) + Lu(t, x) + u(t, x) \partial_t W(t, x) = 0$ are given, where L is a second order elliptic differential operator with random coefficients (dependent on W). To establish such formula the main difficulty is the exponential integrability of some nonlinear integrals, which is proved to be true under some mild conditions on the covariance of W and on the coefficients of L . Along the way, we also obtain an upper bound for increments of stochastic processes on multidimensional rectangles by majorizing measures.

CONTENTS

1. Introduction	2
2. Nonlinear Young integral	5
2.1. Definition	6
2.2. Mapping properties	9
3. Differential equations	11
3.1. Existence and uniqueness	11
3.2. Compositions	15
3.3. Regularity of flow	17
3.4. Transport differential equation	19
4. Feynman-Kac formula - A pathwise approach	21
4.1. Nonlinear Stochastic integral	24
4.2. Feynman-Kac formula I	31
4.3. Feynman-Kac formula II	34
5. Asymptotic growth of Gaussian sample paths	37
5.1. A deterministic inequality	38

Date: April 2014.

2000 *Mathematics Subject Classification.* Primary 60H30; Secondary 60H10, 60H15, 60H07, 60G17.

Key words and phrases. Gaussian random field; sample path property; majorizing measure; nonlinear Young integral; nonlinear Itô-Skorohod integral; transport equation; stochastic parabolic equation; multiplicative noise; Feynman-Kac formula; Malliavin calculus; diffusion process; exponential integrability of the Hölder norm of diffusion process.

Y. Hu is partially supported by a grant from the Simons Foundation #209206 and by a General Research Fund of University of Kansas.

5.2. Majorizing measure	42
5.3. Asymptotic growth	43
Appendix A. Other types of nonlinear stochastic integral	47
A.1. Nonlinear Itô-Skorohod integral	47
A.2. Nonlinear symmetric stochastic integral	53
Appendix B. Estimates for diffusion process	58
Appendix C. Schauder estimates	59
References	65

1. INTRODUCTION

Feynman integral is an important tool in quantum physics. The Feynman-Kac formula is a variant of Feynman integral and plays very important role in the study of (parabolic) partial differential equations (see [20] and [46]). Recently, there have been several successes in extending the Feynman-Kac formula to the following stochastic partial differential equations with noisy (random) potentials on $[0, T] \times \mathbb{R}^d$ (see e.g. [29], [32], and [35]): $\partial_t u(t, x) = \frac{1}{2} \Delta u(t, x) + u(t, x) \partial_t W(t, x)$, where Δ is the Laplacian with respect to spatial variable and $\{\partial_t W(t, x), 0 \leq t \leq T, x \in \mathbb{R}^d\}$ is a Gaussian noise (the derivatives in the sense of Schwartz distribution of a Gaussian field). As indicated in the aforementioned papers, there are three tasks to accomplish for establishing the Feynman-Kac formula. The first one is to give a meaning to the nonlinear stochastic integral $\int_a^b W(ds, x + B_s)$ for a d -dimensional Brownian motion (whose generator is $\frac{1}{2} \Delta$), independent of W . The second one is to establish the exponential integrability of $\int_a^b W(ds, x + B_s)$ and hence the Feynman-Kac expression (which we may call the Feynman-Kac solution) has a rigorous meaning. The final task is to show that the Feynman-Kac expression is indeed a solution to the equation in certain sense. It should be emphasized that the independence between B and W plays crucial role in previous studies.

In many applications, one needs to study more general stochastic partial differential equations. For example, in modeling of the pressure in an oil reservoir in the Norwegian sea with a log normal stochastic permeability one was led to study the stochastic partial differential equation on some bounded domain in \mathbb{R}^d of the form $\operatorname{div}(k(x) \nabla u(x)) = f(x)$, where the permeability $k(x)$ is the (Wick) exponential of white noise, div is the divergence operator, and ∇ is the gradient operator, see [27] and in particular the references therein. Recently, there have been a great amount of research on *uncertainty quantification*. Among the huge literature on this topic let us just mention the books [23], [50], and the references therein. Many different types of stochastic partial differential equations with random coefficients have been studied.

This motivates us to study the Feynman-Kac formula for general stochastic partial differential equations with random coefficients, namely,

$$(1.1) \quad \partial_t u(t, x) + Lu(t, x) + u(t, x) \partial_t W(t, x) = 0,$$

where

$$Lu(t, x) = \frac{1}{2} \sum_{i,j=1}^d a_{ij}(t, x, W) \partial_{x_i x_j}^2 u(t, x) + \sum_{i=1}^d b_i(t, x, W) \partial_{x_i} u(t, x)$$

and for notational simplicity and up to a time change we assume that the terminal condition $u(T, x) = u_T(x)$ is given. The product $u(t, x)\partial_t W(t, x)$ in (1.1) is the ordinary product. If $\sigma(t, x) = (\sigma_{ij}(t, x, W))_{1 \leq i, j \leq d}$ satisfies $a = \sigma\sigma^T$ and if $X_t^{r,x}$ is the solution of the following stochastic differential equation

$$(1.2) \quad dX_t^{r,x} = \sigma(t, X_t^{r,x}, W)\delta B_t + b(t, X_t^{r,x}, W)dt, \quad r \leq t \leq T, \quad X_r^{r,x} = x,$$

then $u(r, x) = \mathbb{E}^B \left\{ u_T(X_T^{r,x}) \exp \left[\int_r^T W(ds, X_s^{r,x}) \right] \right\}$ should be the Feynman-Kac solution to (1.1) with $u(T, x) = u_T(x)$. As indicated above, there are three tasks to complete to justify the above claim. The first task to give a meaning to the nonlinear stochastic integral $\int_r^T W(ds, X_s^{r,x})$ is much more challenging than what has been accomplished before (see for instance [29], [32], and [35]). Although the major focus of the work [32] is to give a meaning to the nonlinear integral $\int_r^T W(ds, X_s^{r,x})$. However, in that paper $X_s^{r,x} = B_s$ is a Brownian motion *independent of* W and then we can consider $X_s^{r,x}$ as “deterministic”. In our current situation since $X_s^{r,x}$ and W are correlated, the nonlinear integral is a true stochastic one. In addition, the noise W may enter to $X_s^{r,x}$ in an anticipative way. Thus, the general stochastic calculus for semimartingales cannot be applied in a straightforward way due to the lack of adaptedness.

If $W(t, x)$ is only continuous in t (without any Hölder continuity in t) but has certain differentiability on x , then we can use semimartingale structure of $X_t^{r,x}$ plus some new techniques developed in Section 4 to define $\int_r^T W(ds, X_s^{r,x})$ and study the corresponding Feynman-Kac solution to (1.1). This result extends the work of [32] in two aspects. One is that the Laplacian is replaced by general second order elliptic operator with general and in particular random coefficients. The other one is that in [32], the Hurst parameter H in time is assumed to be greater than $1/4$, while the result of this paper is applicable to fractional Brownian field whose Hurst parameter H in time can be any number between 0 and 1.

When $W(t, x)$ has certain (Hölder) regularity in time variable, it is natural to see whether one can reduce its regularity in spatial variable x to define $\int_r^T W(ds, X_s^{r,x})$. Having in mind the recent development on rough path analysis and encouraged by the previous success in the case when $X_s^{r,x}$ is the Brownian motion ([29], [32], and [35]), we dedicate ourselves to a systematic study of the nonlinear integral $\int_a^b W(ds, \varphi_s)$, where $W(s, x)$ is a Hölder continuous function on s and x and φ_s is also a Hölder continuous function. Some elementary properties of the integral are obtained as well. These results are presented in Section 2. Let us emphasize that this nonlinear integral $\int_a^b W(ds, \varphi_s)$ is defined in a purely deterministic way. In fact, it is an extension of integration of Young type ([51]).

For Gaussian noise a very important (linear) stochastic integral is the Itô (or Itô-Skorohod) integral. It is also called divergence integral. In probability theory, this integral is a central concept in stochastic analysis. For our stochastic partial differential equation (1.1) it is needed if the product $u(t, x)\partial_t W(t, x)$ there is Wick product. We shall introduce the nonlinear Itô-Skorohod integral $\int_a^b W(ds, \varphi_s)$ (φ_s depends on W) by using Malliavin calculus. This is done in Appendix A. The relation of this integral with other types of integrals is also discussed in this section. Naturally, readers may ask the question to study the Itô-Skorohod type stochastic differential equation $\partial_t u(t, x) + Lu(t, x) + u(t, x) \diamond \partial_t W(t, x) = 0$, where $u(t, x) \diamond W(t, x)$ denotes the Wick product between $u(t, x)$ and $\partial_t W(t, x)$. However, this

seems to be very complex since L depends on W in a sophisticated way and will not be considered in this work.

When $W(t, x)$ is a semimartingale in t for any fixed x and is smooth in x for any fixed t , there has been many studies on stochastic flows which contributes significantly to the study of stochastic partial differential equations (see [37] and the references therein). The important tool there is the nonlinear stochastic integral (with respect to semimartingale) and the corresponding flow. After defining the nonlinear Young integral and motivated by this aspect, we study the pathwise flow associated with time dependent rough vector field $W(t, x)$. That is, we study the differential equation $\varphi_t = x + \int_0^t W(ds, \varphi_s)$ under joint Hölder continuity assumptions of $W(t, x)$. We shall study the flow and other properties of the solution φ_t . This is presented in Section 3. The applications to the transport equation in rough media of the form $\partial_t u(t, x) - \partial_t W(t, x) \nabla u(t, x) = 0$ are also investigated in Subsection 3.4.

After completion of the first task of defining the nonlinear integral another major difficulty (the above mentioned second task) to overcome in the construction of the Feynman-Kac solution is the exponential integrability of $\int_r^T W(ds, X_s^{r,x})$. In the previous work of [29], [32], and [35], this is achieved by showing $\mathbb{E}[u^2(r, x)]$ is finite. If we continue to follow the idea in aforementioned papers, then we are led to show

$$\mathbb{E}^{B, \tilde{B}} \mathbb{E}^W \left\{ u_T(X_T^{r,x}) u_T(\tilde{X}_T^{r,x}) \exp \left[\int_r^T W(ds, X_s^{r,x}) + \int_r^T W(ds, \tilde{X}_s^{r,x}) \right] \right\},$$

is finite, where $\tilde{X}_t^{r,x}$ is the solution to the equation (1.2) with a Brownian motion \tilde{B} , independent of B and W . It seems to us that in our situation, due to the dependence of $X_t^{r,x}$ on W , it is hard to show the above quantity is finite. To get around this difficulty, our strategy is then to show that $u(r, x) = \mathbb{E}^B \left\{ u_T(X_T^{r,x}) \exp \left[\int_r^T W(ds, X_s^{r,x}) \right] \right\}$ is finite for every fixed path of W , assuming some mild pathwise conditions on W (see for instance (4.30)). The third (and the last) task to show that the Feynman-Kac solution is indeed a solution to (1.1) is relatively easier and will be completed by using approximation technique. All these will be done Section 4.

Intentionally, the paper is divided into two parts. The first three chapters can be read without knowledge of probability theory. A single (rough) sample $W(t, x)$ satisfying some joint Hölder continuity and growth conditions is considered. For instance, the (stochastic) partial differential equation (1.1), the nonlinear Young integral (Definition 2.1), and the transport equation (3.24) are considered for every fixed sample path $W(t, x)$. Since $W(t, x)$ is fixed, we also drop the dependence of $a_{ij}(t, x)$ and $b_i(t, x)$ on W throughout the paper. So, the integrals and equations are defined and studied for a (fixed) rough function. The stochastic partial differential equation considered in Section 4 is for a single rough sample path. But Brownian motion is used to represent the solution.

As a probabilist, one may ask whether a stochastic process satisfies the joint Hölder continuity conditions together with the growth conditions assumed throughout the paper. For instance, condition (4.30) in Section 4 requires the paths of W to satisfy

$$(1.3) \quad |W(s, x) - W(s, y) - W(t, x) + W(t, y)| \leq C(1 + |x|^\beta + |y|^\beta) |t - s|^\tau |x - y|^\lambda$$

for all $s, t \in [0, T]$ and $x, y \in \mathbb{R}^d$. We give a partial answer for this problem in Section 5, where an extra assumption $|x - y| \leq \delta$ for a fixed constant δ is imposed. Pathwise boundedness and pathwise regularity (Hölder continuity) have been extensively studied in the literature (see Section 5 for more detailed discussions.) However, estimates similar to (1.3) has not been studied thoroughly. Comparing with the existing literature (e.g. [41], [49]), where estimates for increments over one parameter interval are obtained, the left side of (1.3) is an increment over two parameter rectangle. Difficulties arise because the increments behave differently when the number of parameters get large. For instance, the corresponding entropic volumetric to the left side of (1.3), $d((s, x), (t, y)) = (\mathbb{E}|W(s, x) - W(s, y) - W(t, x) + W(t, y)|^2)^{1/2}$, does not satisfy the triangular inequality. Therefore, classical estimates (such as those appear in [49]) are no longer applicable, new tools are needed to prove (1.3). If in (1.3), x, y are restricted in a compact set, a similar problem has been considered by the authors by extending the Garsia-Rodemich-Rumsey inequality ([30]). Nevertheless, the exact growth rate when x, y get large is not discussed in that paper. Motivated by this requirement, we extend and sharpen our previous work in [30] so that it is applicable to our current situation. Since in many applications, W will be a Gaussian noise, we focus on the case W satisfies normal concentration inequalities to obtain the desirable pathwise property from the covariance structure of the process. As is well-known it is usually hard to obtain properties for each sample path in the theory of stochastic processes. We hope this work will shed some light along this direction.

Notations: We collect here some notations that we will use throughout the entire paper. $A \lesssim B$ means there is a constant C such $A \leq CB$. We represent a vector x in \mathbb{R}^d as a matrix of dimension $d \times 1$, A^T represents the transpose of a matrix A . Sometimes we write x_\bullet for column vector x^T and x^\bullet for the row vector x . We use the Einstein convention on summation over repeated indices. For instance, $b_i c_i$ abbreviates for $\sum_{i=1}^d b_i c_i$.

2. NONLINEAR YOUNG INTEGRAL

Let W and φ be \mathbb{R}^d -valued functions defined on $\mathbb{R} \times \mathbb{R}^d$ and \mathbb{R}^d respectively. We define in the current section the nonlinear Young integration $\int W(ds, \varphi_s)$.

We make the following assumption on the regularity of W

(W) There are constants $\tau, \lambda \in (0, 1]$, $\beta \geq 0$ such that for all $a < b$, the seminorm

$$\begin{aligned}
 & \|W\|_{\beta, \tau, \lambda; a, b} \\
 (2.1) \quad & := \sup_{\substack{a \leq s < t \leq b \\ x, y \in \mathbb{R}^d; x \neq y}} \frac{|W(s, x) - W(t, x) - W(s, y) + W(t, y)|}{(1 + |x| + |y|)^\beta |t - s|^\tau |x - y|^\lambda} \\
 & + \sup_{\substack{a \leq s < t \leq b \\ x \in \mathbb{R}^d}} \frac{|W(s, x) - W(t, x)|}{(1 + |x|)^{\beta + \lambda} |t - s|^\tau} + \sup_{\substack{a \leq t \leq b \\ x, y \in \mathbb{R}^d; x \neq y}} \frac{|W(t, y) - W(t, x)|}{(1 + |x| + |y|)^\beta |x - y|^\lambda},
 \end{aligned}$$

is finite.

About the function φ , we assume

(ϕ) φ is locally Hölder continuous of order $\gamma \in (0, 1]$. That is the seminorm

$$\varphi_{\gamma; a, b} = \sup_{a \leq s < t \leq b} \frac{|\varphi(t) - \varphi(s)|}{|t - s|^\gamma},$$

is finite for every $a < b$.

Throughout the current section, we assume that $\tau + \lambda\gamma > 1$. Among three terms appearing in (2.1), we will pay special attention to the first term. Thus, we denote

$$[W]_{\beta, \tau, \lambda; a, b} = \sup_{\substack{a \leq s < t \leq b \\ x, y \in \mathbb{R}^d; x \neq y}} \frac{|W(s, x) - W(t, x) - W(s, y) + W(t, y)|}{(1 + |x| + |y|)^\beta |t - s|^\tau |x - y|^\lambda}.$$

When $\beta = 0$, then we denote $\|W\|_{\tau, \lambda; a, b} := \|W\|_{0, \tau, \lambda; a, b}$. If a, b are clear in the context, we frequently omit the dependence on a, b . For instance, $\|W\|_{\beta, \tau, \lambda}$ is an abbreviation for $\|W\|_{\beta, \tau, \lambda; a, b}$, $\|\varphi\|_\gamma$ is an abbreviation for $\|\varphi\|_{\gamma; a, b}$ and so on. We shall assume that a and b are finite. It is easy to see that for any $c \in [a, b]$

$$\sup_{a \leq t \leq b} |\varphi(t)| = \sup_{a \leq t \leq b} |\varphi(c) + \varphi(t) - \varphi(c)| \leq |\varphi(c)| + \|\varphi\|_\gamma |b - a|^\gamma < \infty.$$

Thus assumption (ϕ) also implies that

$$\|\varphi\|_{\infty; a, b} := \sup_{a \leq t \leq b} |\varphi(t)| < \infty.$$

For the results presented in this section, the condition (W) can be relaxed to

(W') There are constants $\tau, \lambda \in (0, 1]$, such that for all $a < b$ and compact set K in \mathbb{R}^d , the seminorm

$$\begin{aligned} & \sup_{\substack{a \leq s < t \leq b \\ x, y \in K; x \neq y}} \frac{|W(s, x) - W(t, x) - W(s, y) + W(t, y)|}{|t - s|^\tau |x - y|^\lambda} \\ & + \sup_{\substack{a \leq s < t \leq b \\ x \in K}} \frac{|W(s, x) - W(t, x)|}{|t - s|^\tau} + \sup_{\substack{a \leq t \leq b \\ x, y \in K; x \neq y}} \frac{|W(t, y) - W(t, x)|}{|x - y|^\lambda}, \end{aligned}$$

is finite.

However, the polynomial growth rate is needed in the following sections to solve differential equations.

For later purpose, we denote $C_\beta^{(\tau, \lambda)}(\mathbb{R} \times \mathbb{R}^d)$ (respectively $C_{\text{loc}}^{(\tau, \lambda)}(\mathbb{R} \times \mathbb{R}^d)$) the collection of all functions W satisfying condition (W) (respectively (W')). κ denotes a universal generic constant depending only on λ, τ, α and independent of W, φ and a, b . The value of κ may vary from one occurrence to another.

2.1. Definition. We define the nonlinear integral $\int W(ds, \varphi_s)$ as follows.

Definition 2.1. Let a, b be two fixed real numbers, $a < b$. Let $\pi = \{a = t_0 < t_1 < \dots < t_m = b\}$ be a partition of $[a, b]$ with mesh size $|\pi| = \max_{0 \leq i \leq m-1} |t_{i+1} - t_i|$. The Riemann sum corresponding to π is

$$(2.2) \quad J_\pi = \sum_{i=1}^{m-1} W(t_{i+1}, \varphi_i) - W(t_i, \varphi_i).$$

If the sequence of Riemann sums J_π 's is convergent when $|\pi|$ shrinks to 0, we denote the limit as the nonlinear integral $\int_a^b W(ds, \varphi_s)$.

We observe that in the particular case when $W(t, x) = g(t)x$ for some functions $g : \mathbb{R} \rightarrow \mathbb{R}$, the nonlinear integral $\int_a^b W(ds, \varphi_s)$ defined above, if exists, coincides with the Riemann-Stieltjes integral $\int_a^b \varphi_s dg(s)$. It is well known that if φ and g

are Hölder continuous with exponents α, β respectively and $\alpha + \beta > 1$, then the Riemann-Stieltjes integral $\int_a^b \varphi_s dg(s)$ exists and is called Young integral ([51]).

More generally, for each partition π of an interval $[a, b]$, one can consider the (abstract) Riemann sum

$$(2.3) \quad J_\pi(\mu) = \sum_{i=1}^{m-1} \mu(t_i, t_{i+1})$$

where μ is a function defined on $[a, b]^2$ with values in a Banach space. A sufficient condition for convergence of the limit $\lim_{|\pi| \downarrow 0} J_\pi(\mu)$ is obtained by Gubinelli in [24] via the so-called sewing map. This point of view has important contributions to Lyons' theory of rough paths ([39, 40]). Since we will apply Gubinelli's sewing lemma, we restate the result as follows.

Lemma 2.2 (Sewing lemma). *Let μ be a continuous function on $[a, b]^2$ with values in a Banach space $(B, \|\cdot\|)$ and let $\varepsilon > 0$. Suppose that μ satisfies*

$$\|\mu(s, t) - \mu(s, c) - \mu(c, t)\| \leq K|t - s|^{1+\varepsilon} \quad \forall a \leq s \leq c \leq t \leq b.$$

Then there exists a function $\mathcal{J}\mu(t)$ unique up to an additive constant such that

$$(2.4) \quad \|\mathcal{J}\mu(t) - \mathcal{J}\mu(s) - \mu(s, t)\| \leq K(1 - 2^{-\varepsilon})^{-1}|t - s|^{1+\varepsilon} \quad \forall a \leq s \leq t \leq b.$$

In addition, when $|\pi|$ shrinks to 0, the Riemann sums (2.3) converge to $\mathcal{J}\mu(b) - \mathcal{J}\mu(a)$.

In what follows, we adopt the notation $\mathcal{J}_a^b \mu = \mathcal{J}\mu(b) - \mathcal{J}\mu(a)$. The map $\mu \mapsto \mathcal{J}\mu$ is called the sewing map. The setting of Lemma 2.2 is adopted from [17]. In several occasions, one needs to prove a relation between two or more integrals. The following result provides a simple method for this problem.

Lemma 2.3. *Suppose μ_1 and μ_2 are two functions as in Lemma 2.2. In addition, assume that*

$$|\mu_1(s, t) - \mu_2(s, t)| \leq C|t - s|^{1+\varepsilon'} \quad \forall a \leq s \leq t \leq b$$

for some positive constant ε' . Then $\mathcal{J}\mu_1$ and $\mathcal{J}\mu_2$ are different by an absolute constant. That is $\mathcal{J}_s^t \mu_1 = \mathcal{J}_s^t \mu_2$ for all s, t .

Proof. From Lemma 2.2, $\mathcal{J}(\mu_1 - \mu_2) = \mathcal{J}\mu_1 - \mathcal{J}\mu_2$ and

$$\begin{aligned} |\mathcal{J}_s^t(\mu_1 - \mu_2)| &\lesssim |\mu_1(s, t) - \mu_2(s, t)| + |t - s|^{1+\varepsilon} \\ &\lesssim |t - s|^{1+\varepsilon'} + |t - s|^{1+\varepsilon} \end{aligned}$$

for all s, t . This implies $\mathcal{J}_s^t(\mu_1 - \mu_2) = 0$ for all s, t . \square

Returning to our main objective of the current section, we consider

$$\mu(s, t) = W(t, \varphi_s) - W(s, \varphi_s).$$

Then the condition in Lemma 2.2 is guaranteed by **(W)**, and **(ϕ)**. Indeed, for every $s < c < t$,

$$\begin{aligned} &|\mu(s, t) - \mu(s, c) - \mu(c, t)| \\ &= |W(t, \varphi_s) - W(c, \varphi_s) - W(t, \varphi_c) + W(c, \varphi_c)| \\ &\leq [W]_{\beta, \tau, \lambda} (1 + \|\varphi\|_\infty^\beta) (t - s)^\tau |\varphi_s - \varphi_c|^\lambda \\ &\leq [W]_{\beta, \tau, \lambda} (1 + \|\varphi\|_\infty^\beta) \|\varphi\|_\gamma^\lambda (t - s)^{\tau + \lambda\gamma}. \end{aligned}$$

Hence, by combining the sewing lemma and the previous estimate, we obtain

Proposition 2.4. *Assuming the conditions **(W)**, **(φ)** with $\tau + \lambda\gamma > 1$, the sequence of Riemann sums (2.2) is convergent when $|\pi|$ goes to 0. In other words, the nonlinear integral $\int_a^b W(ds, \varphi_s)$ is well-defined.*

In addition, the following estimate holds

$$(2.5) \quad \left| \int_s^t W(dr, \varphi_r) - W(t, \varphi_c) + W(s, \varphi_c) \right| \leq \kappa \|W\|_{\tau, \lambda; a, b} (1 + \|\varphi\|_\infty^\beta) \|\varphi\|_{\gamma; a, b}^\lambda (t - s)^{\tau + \lambda\gamma}$$

for all $a \leq s \leq c \leq t \leq b$.

Remark 2.5. After the completion of this work, we are brought to the attention of the work [5] (and also [7, 8, 25]), where a similar nonlinear Young integral is studied. The objective of that paper is to define the averaging of the form $\int_0^t f(X_u) du$ for some process X_u and for some irregular function f . The sewing lemma that we follow is from [17], which is after the work of [24].

Remark 2.6. (i) In the particular case when $W(t, x) = g(t)x$, Proposition 2.4 reduces to the existence of the Young integral $\int \varphi_s dg(s)$. Hence, from now on we refer the integral $\int W(ds, \varphi_s)$ as nonlinear Young integral.

(ii) In Proposition 2.4, we can also consider the Riemann sums with right-end points

$$J_\pi^+ = \sum_{i=0}^{m-1} [W(t_{i+1}, \varphi_{t_{i+1}}) - W(t_i, \varphi_{t_{i+1}})].$$

Then the corresponding limit exists and equals to $\int_a^b W(ds, \varphi_s)$. This is a straightforward consequence of Lemma 2.3.

It is evident that

$$\int_s^t W(dr, \varphi_r) = \int_s^c W(dr, \varphi_r) + \int_c^t W(dr, \varphi_r) \quad \forall s < c < t.$$

This together with (2.5) imply easily the following.

Proposition 2.7. *Assume that **(W)** and **(φ)** hold with $\lambda\gamma + \tau > 1$. As a function of t , the indefinite integral $\left\{ \int_a^t W(ds, \varphi_s), a \leq t \leq b \right\}$ is Hölder continuous of exponent τ .*

Fractional calculus is very useful in the study of (linear) Young integral. It leads to some detailed properties of the integral and solution of a differential equation (see [33], [34], and the references therein). It is interesting to extend this approach to nonlinear Young integral. In fact, the authors obtain in [31] the following presentation for the nonlinear Young integral by using fractional calculus. Since this method is not pursued in the current paper, we refer the readers to [31] for further details.

Theorem 2.8. Assume the conditions **(W)** and **(ϕ)** are satisfied. In addition, we suppose that $\lambda\gamma + \tau > 1$. Let $\alpha \in (1 - \tau, \lambda\tau)$. Then the following identity holds

$$(2.6) \quad \begin{aligned} & \int_a^b W(dt, \varphi_t) \\ &= -\frac{1}{\Gamma(\alpha)\Gamma(1-\alpha)} \left\{ \int_a^b \frac{W_{b-}(t, \varphi_t)}{(b-t)^{1-\alpha}(t-a)^\alpha} dt \right. \\ &+ \alpha \int_a^b \int_a^t \frac{W_{b-}(t, \varphi_t) - W_{b-}(t, \varphi_r)}{(b-t)^{1-\alpha}(t-r)^{\alpha+1}} dr dt \\ &+ (1-\alpha) \int_a^b \int_t^b \frac{W(t, \varphi_t) - W(s, \varphi_t)}{(s-t)^{2-\alpha}(t-a)^\alpha} ds dt \\ &\left. + \alpha(1-\alpha) \int_a^b \int_a^t \int_t^b \frac{W(t, \varphi_t) - W(s, \varphi_t) - W(t, \varphi_r) + W(s, \varphi_r)}{(s-t)^{2-\alpha}(t-r)^{\alpha+1}} ds dr dt \right\}, \end{aligned}$$

where $W_{b-}(t, x) = W(t, x) - W(b, x)$.

2.2. Mapping properties. Let μ be a function as in Lemma 2.2. Let us define the quality

$$[\mu]_{1+\varepsilon; I} = \sup_{s, c, t \in I: s < c < t} \frac{|\mu(s, t) - \mu(s, c) - \mu(c, t)|}{|t - s|^{1+\varepsilon}}.$$

In several occasions, given two functions μ_1 and μ_2 such that $[\mu_1]_{1+\varepsilon}$ and $[\mu_2]_{1+\varepsilon}$ are finite, one would like to compare the integrals $\mathcal{J}\mu_1$ and $\mathcal{J}\mu_2$. The following result answers this question.

Lemma 2.9. Let μ_1 and μ_2 be two continuous functions on $[a, b]^2$ such that $[\mu_1]_\alpha$ and $[\mu_2]_\alpha$ are finite for some $\alpha > 1$. Then for every $s, t \in [a, b]$

$$|\mathcal{J}_s^t \mu_1 - \mathcal{J}_s^t \mu_2| \leq |\mu_1(s, t) - \mu_2(s, t)| + (1 - 2^{1-\alpha})^{-1} [\mu_1 - \mu_2]_{\alpha; [s, t]} |t - s|^\alpha$$

Proof. The proof is rather trivial thanks to the linearity nature of Lemma 2.2. Put $\mu = \mu_1 - \mu_2$. Notice that $[\mu]_\alpha \leq [\mu_1]_\alpha + [\mu_2]_\alpha < \infty$. Thus we can apply Lemma 2.2 to μ . The claim follows after observing that $\mathcal{J}\mu = \mathcal{J}\mu_1 - \mathcal{J}\mu_2$. \square

As an application, we study the dependence of the nonlinear Young integration $\int W(ds, \varphi_s)$ with respect to the medium W and the integrand φ .

Proposition 2.10. Let W_1 and W_2 be real valued functions on $\mathbb{R} \times \mathbb{R}^d$ satisfying the condition **(W)**. Let φ be a function in $C^\gamma(\mathbb{R}; \mathbb{R}^d)$ and let $\tau + \lambda\gamma > 1$. Then

$$\begin{aligned} \left| \int_a^b W_1(ds, \varphi_s) - \int_a^b W_2(ds, \varphi_s) \right| &\leq |W_1(b, \varphi_a) - W_1(a, \varphi_a) - W_2(b, \varphi_a) + W_2(a, \varphi_a)| \\ &+ c(\|\varphi\|_\infty) [W_1 - W_2]_{\beta, \tau, \lambda} \|\varphi\|_\gamma |b - a|^{\tau + \lambda\gamma} \end{aligned}$$

Proof. Let $a < c < b$. Put

$$\begin{aligned} \mu_1(a, b) &= W_1(b, \varphi_a) - W_1(a, \varphi_a), \\ \mu_2(a, b) &= W_2(b, \varphi_a) - W_2(a, \varphi_a), \\ \mu &= \mu_1 - \mu_2. \end{aligned}$$

The argument before Proposition 2.4 shows that

$$[\mu]_{\tau+\lambda\gamma} \leq [W_1 - W_2]_{\beta,\tau,\lambda}(1 + \|\varphi\|_\infty^\beta) \|\varphi\|_\gamma.$$

The proposition follows from Lemma 2.9. \square

Proposition 2.11. *Let W be a function on $\mathbb{R} \times \mathbb{R}^d$ satisfying the condition (W). Let φ^1 and φ^2 be two functions in $C^\gamma(\mathbb{R}; \mathbb{R}^d)$ and let $\tau + \lambda\gamma > 1$. Let $\theta \in (0, 1)$ such that $\tau + \theta\lambda\gamma > 1$. Then for any $u < v$*

$$\begin{aligned} \left| \int_u^v W(ds, \varphi_s^1) - \int_u^v W(ds, \varphi_s^2) \right| \\ \leq C_1 [W]_{\beta,\tau,\lambda} \|\varphi^1 - \varphi^2\|_\infty^\lambda |v - u|^\tau \\ + C_2 [W]_{\beta,\tau,\lambda} \|\varphi^1 - \varphi^2\|_\infty^{\lambda(1-\theta)} |v - u|^{\tau+\theta\lambda\gamma}, \end{aligned}$$

where $C_1 = 1 + \|\varphi^1\|_\infty^\beta + \|\varphi^2\|_\infty^\beta$ and $C_2 = 2^{1-\theta} C_1 (\|\varphi^1\|_\gamma^\lambda + \|\varphi^2\|_\gamma^\lambda)^\theta$.

Proof. We put $\mu_1(a, b) = W(b, \varphi_a^1) - W(a, \varphi_a^1)$, $\mu_2(a, b) = W(b, \varphi_a^2) - W(a, \varphi_a^2)$ and $\mu = \mu_1 - \mu_2$. Applying Lemma 2.9, we obtain, for any $\theta \in (0, 1)$ such that $\tau + \theta\lambda\gamma > 1$

$$\begin{aligned} \left| \int_u^v W(ds, \varphi_s^1) - \int_u^v W(ds, \varphi_s^2) \right| \\ \leq |W(v, \varphi_u^1) - W(u, \varphi_u^1) - W(v, \varphi_u^2) + W(u, \varphi_u^2)| \\ + [\mu]_{\tau+\theta\lambda\gamma} |v - u|^{\tau+\theta\lambda\gamma}. \end{aligned}$$

Notice that

$$|W(v, \varphi_u^1) - W(u, \varphi_u^1) - W(v, \varphi_u^2) + W(u, \varphi_u^2)| \leq C_1 [W]_{\beta,\tau,\lambda} |u - v|^\tau \|\varphi^1 - \varphi^2\|_\infty^\lambda.$$

It remains to estimate $[\mu]_{\tau+\theta\lambda\gamma}$. It is obvious that for $i = 1, 2$

$$[\mu_i]_{\tau+\lambda\gamma} \leq [W]_{\beta,\tau,\lambda} (1 + \|\varphi^i\|_\infty^\beta) \|\varphi^i\|_\gamma^\lambda \leq C_1 [W]_{\beta,\tau,\lambda} \|\varphi^i\|_\gamma^\lambda$$

and hence

$$[\mu]_{\tau+\lambda\gamma} \leq [\mu_1]_{\tau+\lambda\gamma} + [\mu_2]_{\tau+\lambda\gamma} \leq C_1 [W]_{\beta,\tau,\lambda} \sum_{i=1}^2 \|\varphi^i\|_\gamma^\lambda.$$

On the other hand

$$\begin{aligned} |\mu(a, b) - \mu(a, c) - \mu(c, b)| \\ \leq |W(b, \varphi_a^1) - W(b, \varphi_a^2) - W(c, \varphi_a^1) + W(c, \varphi_a^2)| \\ + |W(b, \varphi_c^1) - W(b, \varphi_c^2) - W(c, \varphi_c^1) + W(c, \varphi_c^2)| \\ \leq 2C_1 [W]_{\beta,\tau,\lambda} |b - c|^\tau \|\varphi^1 - \varphi^2\|_\infty^\lambda. \end{aligned}$$

Combining the two bounds for μ we get for any $\theta \in (0, 1)$ such that $\tau + \theta\lambda\gamma > 1$,

$$[\mu]_{\tau+\theta\lambda\gamma} \leq C_2 [W]_{\beta,\tau,\lambda} \|\varphi^1 - \varphi^2\|_\infty^{\lambda(1-\theta)}.$$

This completes the proof. \square

Corollary 2.12. *Let I be a nonempty closed, bounded and connected interval. Let t_0 be in I . Assuming condition **(W)** with $\tau + \lambda\gamma > 1$. Then the map*

$$M : C^\gamma(I) \rightarrow C^\tau(I)$$

$$Mx(t) = \int_{t_0}^t W(ds, x_s)$$

is continuous and compact.

Proof. Continuity follows immediately from Proposition 2.11. For compactness, suppose B is a bounded subset of $C^\gamma(I)$. The estimate in Proposition 2.11 implies that $\{Mx\}_{x \in B}$ is bounded in $C^\tau(I)$. By the Arzelà-Ascoli theorem, the set $\{Mx\}_{x \in B}$ is relatively compact in $C^{\tau'}(I)$ for every $\tau' < \tau$. We show that $\{Mx\}_{x \in B}$ is indeed relatively compact in $C^\tau(I)$. More precisely, suppose $\{Mx^n\}$ is a convergent sequence in $M(B)$ in the norm of $C^{\tau'}(I)$, by taking further subsequence, we can assume that the sequence $\{x^n\}$ converges to x in $C^{\gamma'}(I)$, for some $\gamma' < \gamma$ (this is possible since B is bounded). It is sufficient to show that Mx^n converges to Mx in $C^\tau(I)$. To prove this, we choose $\theta \in (0, 1)$ and $\gamma' < \gamma$ such that $\tau + \theta\lambda\gamma' > 1$, and then we apply Proposition 2.11 to obtain

$$\|Mx - Mx^n\|_\tau \leq c\|W\|_{\beta, \tau, \lambda}(\|x - x^n\|_\infty^\lambda + \|x - x^n\|_\infty^{\lambda(1-\theta)}).$$

The constant c depends only on $\|x\|_\infty, \|x\|_{\gamma'}$ and $\|x^n\|_\infty, \|x^n\|_{\gamma'}$ which is uniformly bounded with respect to n . This shows Mx^n converges to Mx in $C^\tau(I)$ and completes the proof. \square

3. DIFFERENTIAL EQUATIONS

Let $W : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ satisfy the condition **(W)** stated at the beginning of Section 2 with $\tau(1 + \lambda) > 1$. In this section we consider the following differential equation

$$(3.1) \quad \varphi_t = \varphi_{t_0} + \int_{t_0}^t W(ds, \varphi_s).$$

We are concerned with the existence, uniqueness, boundedness and the flow property of the solution. We shall also study the dependence of the solution on the initial conditions. Some related results on this direction are also obtained independently by Catellier and Gubinelli [5]. Applications of the results obtained are represented in Subsections 3.3 and 3.4 where we consider a transport equation of the type

$$u(dt, x) = \nabla u(t, x)W(dt, x).$$

Literature on transport equations is vast and mostly focuses on irregularity of the spatial variables of the vector field (see for instance [13] for Sobolev vector fields, [2] for BV vector fields and [3] for Besov vector fields). In the case W being a semi-martingale, the above equation is treated in [37]. It appears to be new in the context of nonlinear Young integration.

3.1. Existence and uniqueness.

Theorem 3.1 (Existence). *Suppose that W satisfies the assumption **(W)** with $\tau(1 + \lambda) > 1$ and $\beta + \lambda \leq 1$. Then the equation (3.1) has a solution in the space of*

Hölder continuous functions $C^\tau([t_0 - T, t_0 + T])$ for any $T > 0$. Moreover, if φ is a solution in $C^\tau([t_0 - T, t_0 + T])$, then

(3.2)

$$\sup_{t \in [t_0 - T, t_0 + T]} |\varphi_t| + \sup_{t_0 - T \leq s < t \leq t_0 + T} \frac{|\varphi_t - \varphi_s|}{|t - s|^\tau} \leq C_{\tau, \lambda, T} e^{\kappa_{\tau, \lambda, T} \|W\|_{\tau, \lambda}^{\frac{1 - \tau + \tau \lambda}{\tau \lambda}}} (1 \vee |\varphi_{t_0}|),$$

where the constant $\kappa_{\tau, \lambda, T}$ and $C_{\tau, \lambda, T}$ depend only on λ , τ and T .

Proof. Fix $T > 0$, we denote $\|W\| = \|W\|_{\beta, \tau, \lambda; [t_0 - T, t_0 + T]}$. We define a mapping M acting on $C^\tau([t_0 - T, t_0 + T])$ as follows

$$Mx = x_0 + \int_{t_0}^{\cdot} W(ds, x_s), \quad \forall x \in C^\tau([t_0 - T, t_0 + T]).$$

We shall verify that M satisfies the hypothesis of Leray-Schauder theorem (see [22, Theorem 11.3]).

Step 1. M is well-defined, continuous and compact. This immediately follows from Corollary 2.12.

Step 2. Now we explain that the set $\{x \in C^\tau([t_0 - T, t_0 + T]) : x = \sigma Mx, 0 \leq \sigma \leq 1\}$ is bounded. Let x satisfy $x = \sigma Mx$ for some $\sigma \in [0, 1]$. By definition of M , we see $x = \sigma Mx$ can be written as

$$x_b - x_a = \sigma \int_a^b W(ds, x_s).$$

From (2.5), it follows that for any $a, b \in [t_0 - T, t_0 + T]$, we have

$$\begin{aligned} |x_b - x_a| &= \sigma \left| \int_a^b W(ds, x_s) \right| \\ &\leq \sigma \|W\| (1 + \|x\|_{\infty; a, b}^\beta) \|x\|_{\infty; a, b}^\lambda (b - a)^\tau \\ &\quad + \sigma \kappa \|W\| (1 + \|x\|_{\infty; a, b}^\beta) \|x\|_{\tau; a, b}^\lambda |b - a|^{\tau + \lambda \tau}. \end{aligned}$$

Since $\sigma \leq 1$, this yields

$$\|x\|_{\tau; a, b} \leq \|W\| (1 + \|x\|_{\infty; a, b}^\beta) \|x\|_{\infty; a, b}^\lambda + \kappa \|W\| (1 + \|x\|_{\infty; a, b}^\beta) \|x\|_{\tau; a, b}^\lambda |b - a|^{\lambda \tau},$$

for every a, b in $[t_0, t_0 + T]$ with $a < b$. We emphasize that the constant κ appears in the previous inequality is independent of σ . An application of Young inequality gives

$$\|x\|_{\infty; a, b}^\beta \|x\|_{\tau; a, b}^\lambda \leq \|x\|_{\infty; a, b}^{\beta + \lambda} + \|x\|_{\tau; a, b}^{\beta + \lambda}.$$

Thus

$$\begin{aligned} \|x\|_{\tau; a, b} &\leq \|W\| (\|x\|_{\infty; a, b}^\lambda + \|x\|_{\infty; a, b}^{\beta + \lambda}) + \kappa \|W\| \|x\|_{\infty; a, b}^{\beta + \lambda} |b - a|^{\lambda \tau} \\ &\quad + \kappa \|W\| (\|x\|_{\tau; a, b}^\lambda + \|x\|_{\tau; a, b}^{\beta + \lambda}) |b - a|^{\lambda \tau}. \end{aligned}$$

Applying the inequality $z^\theta \leq 1 \vee z$ ($\theta \in [0, 1]$ and $z \geq 0$), we obtain

$$\|x\|_{\tau; a, b} \leq \|W\| (2 + \kappa |b - a|^{\lambda \tau}) (1 \vee \|x\|_{\infty; a, b}) + \kappa \|W\| (1 \vee \|x\|_{\tau; a, b}) |b - a|^{\lambda \tau}.$$

We further use

$$\|x\|_{\infty; a, b} \leq |x_a| + \|x\|_{\tau; a, b} |b - a|^\tau$$

to obtain

$$(3.3) \quad \|x\|_{\tau; a, b} \leq A \|W\| (1 \vee |x_a|) + A \|W\| (1 \vee \|x\|_{\tau; a, b}) |b - a|^{\lambda \tau},$$

where A is a constant depending only on τ, λ and T . Let Δ be a positive number such that

$$(3.4) \quad A\|W\|\Delta^{\tau\lambda} = \frac{1}{2}.$$

If $|b - a| \leq \Delta$, then from (3.3)

$$(3.5) \quad \|x\|_{\tau;a,b} \leq 2A\|W\|(1 \vee |x_a|).$$

Hence, we obtain

$$(3.6) \quad (1 \vee \|x\|_{\infty,a,b}) \leq (2A\|W\|\Delta^{\tau} + 1)(1 \vee |x_a|).$$

Divide the interval $[t_0, t_0 + T]$ into $n = [T/\Delta] + 1$ subintervals of length less or equal than Δ . Applying the inequality (3.6) on the intervals $[t_0, t_0 + \Delta], [t_0 + \Delta, \dots], [t_0 + (n-1)\Delta, t_0 + n\Delta \wedge T]$, recursively, we obtain

$$(3.7) \quad (1 \vee \|x\|_{\infty,t_0,t_0+T}) \leq (2A\|W\|\Delta^{\tau} + 1)^n (1 \vee |x_{t_0}|).$$

We can also assume that $\Delta \leq T$. Thus $n \leq 2T/\Delta$. We use the bound $2A\|W\|\Delta^{\tau} + 1 \leq \exp(2A\|W\|\Delta^{\tau})$. Then (3.7) yields

$$(1 \vee \|x\|_{\infty,t_0,t_0+T}) \leq \exp(2A\|W\|\Delta^{\tau} \frac{2T}{\Delta})(1 \vee |x_{t_0}|).$$

Using (3.4), namely,

$$\Delta = (2A\|W\|)^{-\frac{1}{\tau\lambda}},$$

we have

$$(1 \vee \|x\|_{\infty,t_0,t_0+T}) \leq e^{T(2A\|W\|)^{\frac{1-\tau+\tau\lambda}{\tau\lambda}}}(1 \vee |x_{t_0}|),$$

where $C_{\tau,\lambda}$ and $\kappa_{\tau,\lambda}$ are uniformly bounded in $\sigma \in [0, 1]$. The argument goes similarly on the other interval $[t_0 - T, t_0]$. Thus

$$(3.8) \quad (1 \vee \|x\|_{\infty,t_0-T,t_0+T}) \leq e^{T(2A\|W\|)^{\frac{1-\tau+\tau\lambda}{\tau\lambda}}}(1 \vee |x_{t_0}|).$$

Together with the estimate (3.5), this inequality (3.8) implies that the set

$$\{x \in C^{\tau}([t_0 - T, t_0 + T]) : x = \sigma Lx, 0 \leq \sigma \leq 1\}$$

is bounded in $C^{\tau}([t_0 - T, t_0 + T])$.

Step 3. Applying Leray-Schauder theorem, we see that the equation (3.1) has a solution $\{\varphi_t, t \in [t_0 - T, t_0 + T]\}$ in $C^{\tau}([t_0 - T, t_0 + T])$ for every T . The estimate (3.2) comes from (3.8) together with (3.5). \square

Next, we study some stability result. In particular, we want to know how the solution depends on the initial condition x_{t_0} .

Theorem 3.2. *Let the condition (W) be satisfied with $\tau + \tau\lambda > 1$. In addition, we assume that $W(t, x)$ is differentiable with respect to x for every t and the spatial gradient matrix of W is denoted by $\nabla W(t, x) = \left(\frac{\partial W_i(t, x)}{\partial x_j} \right)_{1 \leq i, j \leq d}$. Suppose*

$$\begin{aligned} \|\nabla W\|_{\tau,\lambda;[t_0-T,t_0+T] \times K} &:= \sup_{\substack{t_0-T \leq s < t \leq t_0+T \\ x \in K}} \frac{|\nabla W(t, x) - \nabla W(s, x)|}{|t - s|^{\tau}} \\ &+ \sup_{\substack{t_0-T \leq s < t \leq t_0+T \\ x, y \in K, x \neq y}} \frac{|\nabla W(t, x) - \nabla W(s, x) - \nabla W(t, y) + \nabla W(s, y)|}{|t - s|^{\tau} |x - y|^{\lambda}} \end{aligned}$$

is finite for all compact set K in \mathbb{R}^d . Let x_t and y_t be two solutions in $C^\tau([t_0 - T, t_0 + T])$ to the integral equation (3.1) with initial conditions x_0 and y_0 respectively. Then the following estimate holds

$$(3.9) \quad \sup_{t \in [t_0 - T, t_0 + T]} |x_t - y_t| \leq 2^{\kappa T A^{\frac{1}{\tau}}} |x_0 - y_0|,$$

where A is a constant depending on ∇W , x, y and T (precise formula is given in (3.10) below).

Proof. We put $R = \max\{\|x\|_{\infty; [t_0 - T, t_0 + T]}, \|y\|_{\infty; [t_0 - T, t_0 + T]}\}$, $K = \{x \in \mathbb{R}^d : |x| \leq R\}$ and $\|\nabla W\| = \|\nabla W\|_{\tau, \lambda; [t_0 - T, t_0 + T] \times K}$. We also denote $z_t = x_t - y_t$, $\rho_\tau = (\|x\|_\tau + \|y\|_\tau)^\lambda$ and $\eta_t = \eta x_t + (1 - \eta)y_t$ for each $\eta \in (0, 1)$. For every s, t and x , we use the notation $W([s, t], x) = W(t, x) - W(s, x)$.

We shall obtain estimate for z in $C([t_0 - T, t_0 + T])$. Fix $a < b$ in $[t_0 - T, t_0 + T]$. We then write

$$z_b - z_a = \int_a^b W(ds, x_s) - \int_a^b W(ds, y_s) = \mathcal{J}_a^b \mu,$$

where μ is the function

$$\mu(s, t) = W([s, t], x_s) - W([s, t], y_s) = \int_0^1 \nabla W([s, t], \eta_s) z_s d\eta.$$

For every $s \leq c \leq t$ in $[a, b]$, we can write

$$\begin{aligned} \mu(s, t) - \mu(s, c) - \mu(c, t) \\ = \int_0^1 ([\nabla W([c, t], \eta_s) - \nabla W([c, t], \eta_c)] z_s + \nabla W([c, t], \eta_c)(z_t - z_c)) d\eta. \end{aligned}$$

We note that $|\eta_t - \eta_s|^\lambda = |\eta(x_t - x_s) + (1 - \eta)(y_t - y_s)|^\lambda \leq \rho_\tau |u - v|^{\tau\lambda}$. It follows that

$$[\mu]_{\tau(1+\lambda); [a, b]} \leq \|\nabla W\|(\rho_\tau \|z\|_{\infty; a, b} + |b - a|^{\tau(1-\lambda)} \|z\|_{\tau; a, b}).$$

On the other hand, it is obvious that $|\mu(a, b)| \leq \|\nabla W\| |b - a|^\tau \|z\|_{\infty; a, b}$. Hence, the estimate (2.4) implies

$$|z_b - z_a| \leq \|\nabla W\| |b - a|^\tau \|z\|_{\infty; a, b} + \kappa \|\nabla W\| |b - a|^{\tau + \lambda\tau} (\rho_\tau \|z\|_{\infty; a, b} + |b - a|^{\tau(1-\lambda)} \|z\|_{\tau; a, b}).$$

In other words,

$$\|z\|_{\tau; a, b} \leq A \|z\|_{\infty; a, b} + A \|z\|_{\tau; a, b} (b - a)^\tau,$$

where

$$(3.10) \quad A = \kappa \|\nabla W\| [1 + \rho_\tau T^{\lambda\tau}].$$

Therefore, using the bound $\|z\|_{\infty; a, b} \leq |z_a| + \|z\|_{\tau; a, b}$ one gets

$$(3.11) \quad \|z\|_{\tau; a, b} \leq A |z_a| + 2A \|z\|_{\tau; a, b} (b - a)^\tau.$$

Now we shall use the above inequality to show our theorem. Choose a, b such that

$$|b - a| \leq \Delta = \left(\frac{1}{4A} \right)^{\frac{1}{\tau}}.$$

Then inequality (3.11) implies $\|z\|_{\tau,a,b} \leq 2A|z_a|$ for all $a < b$. By the definition of the Hölder norm, we see that if $|b - a| \leq \Delta$, then

$$\begin{aligned} \|z\|_{\infty,a,b} &\leq |z_a| + \|z\|_{\tau,a,b}(b-a)^\tau \\ &\leq |z_a| + 2A|z_a|\Delta^\tau \\ &\leq 2|z_a|. \end{aligned}$$

Divide the interval $[t_0, t_0 + T]$ into $n = [T/\Delta] + 1$ subintervals of length less or equal than Δ . Applying the previous inequality on the intervals $[t_0, t_0 + \Delta]$, $[t_0 + \Delta, t_0 + 2\Delta]$, \dots , $[t_0 + (n-1)\Delta, t_0 + n\Delta \wedge T]$, recursively, we obtain

$$\|z\|_{\infty,t_0,t_0+T} \leq 2^n |z_{t_0}|.$$

We can assume $\Delta \leq T$. Thus

$$n = [T/\Delta] + 1 \leq \frac{2T}{\Delta} = 2T(4A)^\frac{1}{\tau}.$$

This implies

$$\|z\|_{\infty,t_0,t_0+T} \leq 2^{2^{1+2/\tau}TA^\frac{1}{\tau}} |z_{t_0}|.$$

which yields the bound (3.9) on the interval $[t_0, t_0 + T]$. Estimates on $[t_0 - T, t_0]$ are analogous. \square

An immediate consequence of the theorem is the following uniqueness result.

Corollary 3.3. *Under the hypothesis of Theorem 3.2 the equation (3.1) has a unique solution.*

3.2. Compositions. Given a function $G : \mathbb{R}^2 \rightarrow \mathbb{R}^d$, we may define the Riemann-Stieltjes integral $\int_a^b G(ds, s)$ as the limit of Riemann sums

$$\sum_i G(t_i, t_{i-1}) - G(t_{i-1}, t_{i-1}).$$

The sewing lemma (Lemma 2.2) gives a sufficient condition so that the aforementioned limit exists, namely G satisfies

$$|G(s, s) - G(s, t) - G(t, s) + G(t, t)| \lesssim |t - s|^{1+\varepsilon}$$

for some $\varepsilon > 0$. In such case, Lemma 2.3 also allows one to choose Riemann sums with right-end points. In other words, the Riemann sums with right-end points

$$\sum_i G(t_i, t_i) - G(t_{i-1}, t_i)$$

also converges to the Riemann-Stieltjes integral $\int_a^b G(ds, s)$. In what follows, all integrals are understood as Riemann-Stieltjes integration, except for a few occasions, which we will indicate. The following result can be regarded as Itô formula or chain rule for compositions of functions in the context of nonlinear Young integration.

Theorem 3.4. *Let F be a function in $C_{\text{loc}}^{(\tau_F, \lambda_F)}(\mathbb{R} \times \mathbb{R}^d)$ (i.e. F satisfies the condition (W') with τ_F and λ_F), g and x be Hölder continuous functions with exponents τ_g and τ respectively. We suppose that $\tau_F + \lambda_F \tau > 1$ and $\tau_g + \tau_F > 1$. The following integration by parts formula holds*

$$(3.12) \quad \int_0^T g(t) dF(t, x_t) = \int_0^T g(t) F(dt, x_t) + \int_0^T g(t) F(t, dx_t).$$

In particular, suppose that F belongs to $C_{\text{loc}}^{\tau_F}(\mathbb{R}; C_{\text{loc}}^{1+\lambda_F}(\mathbb{R}^d))$, x is of the form $x_t = \int_a^t W(ds, \phi_s)$, where W satisfy the condition **(W')** with τ and λ , ϕ satisfy **(\phi)** with γ , $\tau + \lambda\gamma > 1$ and $\tau\lambda_F + \tau > 1$. Then (3.12) becomes

$$(3.13) \quad \int_0^T g(t) dF(t, x_t) = \int_0^T g(t) F(dt, x_t) + \int_0^T g(t) (\nabla F)(t, x_t) W(dt, \phi_t).$$

An important consequence of (3.13) is when g is a constant function

$$(3.14) \quad F(b, x_b) - F(a, x_a) = \int_a^b F(dt, x_t) + \int_a^b (\nabla F)(t, x_t) W(dt, \phi_t).$$

Proof. We choose a compact set K such that K contains $\{x_t, 0 \leq t \leq T\}$ and denote $\|F\| = \|F\|_{\tau_F, \lambda_F; [0, T] \times K}$. We put

$$\begin{aligned} \mu(a, b) &= g(b)F(b, x_b) - g(b)F(a, x_b), \\ \nu(a, b) &= g(a)F(a, x_b) - g(a)F(a, x_a), \\ \vartheta(a, b) &= g(a)F(b, x_b) - g(a)F(a, x_a). \end{aligned}$$

For every $a < c < b$, we have

$$\begin{aligned} & |\mu(a, b) - \mu(a, c) - \mu(c, a)| \\ &= | -g(b)F(a, x_b) - g(c)F(c, x_c) + g(c)F(a, x_c) + g(b)F(c, x_b) | \\ &\leq |g(c)| |F(a, x_b) - F(c, x_c) + F(a, x_c) + F(c, x_b)| \\ &\quad + |g(c) - g(b)| |F(c, x_c) - F(a, x_c)| \\ &\leq \|g\|_{\infty} \|F\| \|x\|_{\tau}^{\lambda_F} |b - a|^{\tau_F + \lambda_F \tau} + \|g\|_{\tau_g} \|F\| |b - a|^{\tau_g + \tau_F}, \end{aligned}$$

and

$$\begin{aligned} & |\nu(a, b) - \nu(a, c) - \nu(c, a)| \\ &= |g(a)F(a, x_b) - g(a)F(a, x_c) - g(c)F(c, x_b) + g(c)F(c, x_c)| \\ &\leq |g(c)| |F(a, x_b) - F(a, x_c) - F(c, x_b) + F(c, x_c)| \\ &\quad + |g(a) - g(c)| |F(a, x_b) - F(a, x_c)| \\ &\lesssim \|g\|_{\infty} \|F\| |b - a|^{\tau_F + \lambda_F \tau} + \|g\|_{\tau_g} \|F\| \|x\|_{\tau}^{\lambda_F} |b - a|^{\tau_g + \lambda_F \tau}. \end{aligned}$$

Hence, from Lemmas 2.2 and 2.3, $\mathcal{J}_0^T \mu = \int_0^T g(t) F(dt, x_t)$ and $\mathcal{J}_0^T \nu = g(t) F(t, dx_t)$. On the other hand,

$$\begin{aligned} & |\vartheta(a, b) - \mu(a, b) - \nu(a, b)| \\ &= |[g(a) - g(b)][F(b, x_b) - F(a, x_b)]| \leq \|g\|_{\tau_g} \|F\| |b - a|^{\tau_g + \tau_F}. \end{aligned}$$

This together with Lemma 2.3 implies (3.12).

To prove (3.13), it suffices to show

$$(3.15) \quad \int_0^T g(t) F(t, dx_t) = \int_0^T g(t) (\nabla F)(t, x_t) W(dt, \phi_t).$$

We put

$$\tilde{\nu}(a, b) = g(a) \nabla F(a, x_a) [W(b, \phi_a) - W(a, \phi_a)].$$

Then we write

$$\begin{aligned}\nu(a, b) &= g(a) \int_0^1 \nabla F(a, \eta x_a + (1 - \eta)x_b) d\eta (x_a - x_b) \\ &= g(a) \int_0^1 \nabla F(a, \eta x_a + (1 - \eta)x_b) d\eta \int_a^b W(ds, \phi_s).\end{aligned}$$

Using the estimate (2.5), we obtain

$$\begin{aligned}|\nu(a, b) - \tilde{\nu}(a, b)| &\leq |g(a) \int_0^1 [\nabla F(a, \eta x_a + (1 - \eta)x_b) - \nabla F(a, x_a)] d\eta \int_a^b W(ds, \phi_s)| \\ &\quad + |g(a) \nabla F(a, x_a) [\int_a^b W(ds, \phi_s) - W(b, \phi_b) + W(a, \phi_a)]| \\ &\lesssim |b - a|^{\lambda_F \tau + \tau} + |b - a|^{\tau + \lambda \gamma}.\end{aligned}$$

Identity (3.15) follows from Lemma 2.3 and the previous estimate. \square

3.3. Regularity of flow. In the rest of the current section, we assume the hypothesis of Theorem 3.2. This assumption guarantees that $\varphi(t, x)$, the solution to

$$\varphi(t, x) = x + \int_0^t W(ds, \varphi(s, x))$$

is unique. Moreover, by the result in Subsection 3.1, for fixed t , $\varphi(t, \cdot)$ is an automorphism on \mathbb{R}^d , its inverse is $\varphi(t, \cdot)^{-1} = \varphi(-t, \cdot)$. Hence, the family $\{\varphi(t, \cdot) : t \in \mathbb{R}\}$ forms a flow of homeomorphism, i.e. it satisfies the following properties:

- $\varphi(t + s, \cdot) = \varphi(t, \varphi(s, \cdot))$ holds for all s, t ,
- $\varphi(0, \cdot)$ is the identity map,
- the map $\varphi(t, \cdot) : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a homeomorphism for all t .

Moreover, one can show that $\varphi(t, \cdot)$ is indeed a diffeomorphism.

Theorem 3.5. *Assume the hypothesis of Theorem 3.2. For any t in \mathbb{R} , the map $\varphi(t, \cdot)$ is a diffeomorphism. The following conclusions hold*

- (i) *The gradient of φ_t at x , denoted by $\nabla \varphi(t, x) = \{\partial_j \varphi^i(t, x)\}_{i,j}$ satisfies the equation*

$$(3.16) \quad \partial_i \varphi^\bullet(t, x) = \delta_{\bullet i} + \int_0^t \partial_k W^\bullet(ds, \varphi(s, x)) \partial_k \varphi^i(s, x)$$

where δ_{ij} is the Kronecker symbol. Equation (3.16) can be written in short

$$\nabla \varphi(t, x) = I_d + \int_0^t \nabla W(ds, \varphi(s, x)) \nabla \varphi(s, x).$$

- (ii) *For every t and x , the matrix $\nabla \varphi(t, x)$ is invertible and its inverse $M(t, x) = [\nabla \varphi(t, x)]^{-1}$ satisfies the equation*

$$(3.17) \quad M(t, x)^{j\bullet} = \delta_{j\bullet} - \int_0^t M(s, x)^{jk} \partial_\bullet W^k(ds, \varphi(s, x))$$

or in short

$$M(t, x) = I_d - \int_0^t M(s, x) \nabla W(ds, \varphi(s, x)).$$

(iii) φ is jointly Hölder continuous of order $(\tau, 1)$. That is

$$(3.18) \quad |\varphi(s, x) - \varphi(s, y) - \varphi(t, x) + \varphi(t, y)| \lesssim |t - s|^\tau |x - y|$$

(iv) Let $J(t, x)$ denote the determinant of $\nabla \varphi(t, x)$. Then J satisfies the following scalar linear equation

$$(3.19) \quad J(t, x) = 1 + \int_0^t J(s, x) \operatorname{Div}(W(ds, \varphi(s, x))).$$

(v) The flow $\varphi(t, x)$ is a Lagrangian flow, namely there exists a constant L such that

$$(3.20) \quad \mathcal{L}^d(\varphi(t, \cdot)^{-1}(A)) \leq L \mathcal{L}^d(A) \quad \text{for every Borel set } A \subseteq \mathbb{R}^d$$

where \mathcal{L}^d is the Lebesgue measure on \mathbb{R}^d .

Proof. Let e be a unit vector in \mathbb{R}^d . For each h in \mathbb{R} , we denote

$$\eta_t^h = \frac{1}{h}(\varphi(t, x + he) - \varphi(t, x)).$$

To prove (i), it is sufficient to show that for every sequence h_n converging to 0, there is a subsequence h_{n_k} such that $\eta^{h_{n_k}}$ converges to the solution of the following equation

$$(3.21) \quad \eta_t = e + \int_0^t \nabla W(ds, \varphi(s, x)) \eta_s.$$

We remark that the equation (3.21) is linear and the existence and uniqueness of solution in $C^\tau(\mathbb{R})$ follows from our method discussed in Subsection 3.1. From Theorem 3.2 we see that

$$\|\eta^h\|_{\tau; K} \leq \kappa_K$$

uniformly in h for every compact interval K in \mathbb{R} . Hence, by the Arzelà-Ascoli theorem, there is a subsequence, still denoted by h_n such that η^{h_n} converges to η in $C^{\tau'}(K)$ for any arbitrary $\tau' < \tau$. On the other hand, we notice that η^h satisfies

$$(3.22) \quad \eta_t^h = e + \int_0^1 d\tau \int_0^t \nabla W(ds, \tau \varphi(s, x + he) - (1 - \tau) \varphi(s, x)) \eta_s^h.$$

Passing through the limit $h_n \rightarrow 0$, we see that η satisfies the equation (3.21) and then (i) follows. Assertion (iii) is a consequence of the estimate (3.9) in Theorem 3.2. In fact,

$$\begin{aligned} |\varphi(s, x) - \varphi(s, y) - \varphi(t, x) + \varphi(t, y)| &\leq \|\varphi(\cdot, x) - \varphi(\cdot, y)\|_{\tau; [s, t]} |t - s|^\tau \\ &\lesssim |t - s|^\tau |x - y|. \end{aligned}$$

Assertion (iv) follows from the Itô formula (3.14) applied to $J(t, x) = \det(\nabla \varphi(t, x))$ and the Jacobi's formula

$$d \det(M) = \det(M) \operatorname{tr}(M^{-1} dM).$$

To prove (v), we notice that the equation (3.19) can be solved explicitly thanks to (3.14)

$$(3.23) \quad J(t, x) = \exp \int_0^t \operatorname{Div}(W(ds, \varphi(s, x))).$$

Therefore, from (2.5), we obtain

$$|J(t, x)^{-1}| \leq e^{\kappa |t|^\tau}.$$

Together with the area formula

$$\mathcal{L}^d(\varphi(t, \cdot)^{-1}(A)) = \int_{\varphi(-t, A)} dx = \int_A |\det(\nabla \varphi)(-t, x)| dx$$

this estimate implies (3.20). \square

3.4. Transport differential equation. As an application of the above Itô formula (3.13) and flow property (Theorem 3.5), we study the following transport differential equation in *Hölder media*. Specifically, let $W : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ satisfy the conditions in Theorem 3.2. Consider the following first order partial differential equations (transport equation in Hölder media W)

$$(3.24) \quad \frac{\partial}{\partial t} u(t, x) + \left(\frac{\partial}{\partial t} W(t, x) \right) \cdot \nabla u(t, x) = 0.$$

Here ∇ is the gradient operator (with respect to spatial variables). Since W is only Hölder continuous in time, the equation (3.24) is only formal. We can however define solutions in integral form. More precisely, a continuous function $u : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}$ is called a solution to (3.24) with the initial condition $u(0, x) = h(x)$ if it is differentiable with respect to $x \in \mathbb{R}^d$ and the following equation holds.

$$(3.25) \quad u(t, x) = h(x) - \int_0^t \nabla u(s, x) W(ds, x) \quad \forall t \geq 0, x \in \mathbb{R}^d.$$

Theorem 3.6. *Assuming W satisfies the conditions in Theorem 3.2. Let h be a function in $C_{\text{loc}}^{1+\lambda_0}(\mathbb{R}^d)$ where λ_0 satisfies $(1 + \lambda_0)\tau > 1$. Let $\varphi(t, x)$ be the unique solution to*

$$\varphi(t, x) = x + \int_0^t W(ds, \varphi(s, x)), \quad \forall t \geq 0.$$

Let $\psi(t, x)$ be the inverse of φ as a function $x \in \mathbb{R}^d$ to \mathbb{R}^d . Namely, $\varphi(t, \psi(t, x)) = x$ for all $t \geq 0, x \in \mathbb{R}^d$. Then the function u defined by

$$u(t, x) = h(\psi(t, x))$$

is a solution to the above transport equation.

Proof. From Theorem 3.5 such $\psi(t, x)$ exists and both $\varphi(t, x)$ and $\psi(t, x)$ are differentiable with respect to x . Differentiate $\varphi(t, \psi(t, x)) = x$ with respect to x and we see that

$$(\nabla \varphi)(t, \psi(t, x)) \nabla \psi(t, x) = I,$$

or

$$(\nabla \psi(t, x))^{-1} = (\nabla \varphi)(t, \psi(t, x)).$$

Let $\rho(r) = \varphi(r, \psi(r, x))$, $0 \leq r < \infty$. Thanks to Theorem 3.5(iii), Itô formula (3.13) is applicable. More precisely, for any C^τ -function $g(r)$, we have

$$\int_0^t g(r) d\rho(r) = \int_0^t g(r) \varphi(dr, \psi(r, x)) + \int_0^t g(r) (\nabla \varphi)(r, \psi(r, x)) \psi(dr, x).$$

Since $\rho(r) = x$, we have $d\rho(r) = 0$. Thus

$$(3.26) \quad \int_0^t g(r) (\nabla \varphi)(r, \psi(r, x)) \psi(dr, x) = - \int_0^t g(r) \varphi(dr, \psi(r, x)).$$

Now the Itô formula (3.14) applied to $h(\psi(t, x))$ yields

$$\begin{aligned}
u(t, x) &= h(\psi(t, x)) = h(x) + \int_0^t (\nabla h)(\psi(r, x)) \psi(dr, x) \\
&= h(x) + \int_0^t \nabla [h(\psi(r, x))] (\nabla \psi(r, x))^{-1} \psi(dr, x) \\
&= h(x) + \int_0^t \nabla u(r, x) (\nabla \psi(r, x))^{-1} \psi(dr, x) \\
&= h(x) + \int_0^t \nabla u(r, x) (\nabla \varphi)(r, \psi(r, x)) \psi(dr, x).
\end{aligned}$$

Using the equation (3.26) for $g(r) = \nabla u(r, x)$, we have

$$\begin{aligned}
u(t, x) &= h(x) - \int_0^t \nabla u(r, x) \varphi(dr, \psi(r, x)) \\
&= h(x) - \int_0^t \nabla u(r, x) W(dr, \varphi(r, \psi(r, x))) \\
&= h(x) - \int_0^t \nabla u(r, x) W(dr, x).
\end{aligned}$$

This completes the proof of the theorem. \square

We also have the following uniqueness result.

Theorem 3.7. *Assuming W satisfies the conditions in Theorem 3.2. Let λ_0 be in $(0, 1]$ such that $(\lambda_0 + 1)\tau > 1$. Equation (3.25) has unique solution in the class $C_{\text{loc}}^{(\tau, \lambda_0)}(\mathbb{R} \times \mathbb{R}^d)$. More precisely, suppose u belongs to $C_{\text{loc}}^{(\tau, \lambda_0)}(\mathbb{R} \times \mathbb{R}^d)$ and satisfies (3.25), then u is uniquely defined by the relation $u(t, x) = h(\psi(t, x))$, where φ and ψ are the functions defined in Theorem 3.6.*

Proof. Let u be a solution to (3.25). Applying Itô formula (3.14) for the function $u(t, \varphi(t, x))$ we have

$$u(t, \varphi(t, x)) - h(x) = \int_0^t u(ds, \varphi(s, x)) + \int_0^t \nabla u(s, \varphi(s, x)) W(ds, \varphi(s, x)).$$

It suffices to show the right hand side vanishes. In other words the following relation between the two nonlinear Young integrals holds

$$(3.27) \quad \int_0^t u(ds, \varphi(s, x)) = - \int_0^t \nabla u(s, \varphi(s, x)) W(ds, \varphi(s, x)).$$

For clarity, we will omit x in the notations. We put

$$\begin{aligned}
\mu_1(a, b) &= u(b, \varphi_a) - u(a, \varphi_a), \\
\mu_2(a, b) &= \nabla u(a, \varphi_a)[W(b, \varphi_a) - W(a, \varphi_a)].
\end{aligned}$$

Since u satisfies the equation (3.25), we can write

$$\mu_1(a, b) = - \int_a^b \nabla u(s, \varphi_a) W(ds, \varphi_a).$$

Thus

$$\begin{aligned} \mu_1(a, b) + \mu_2(a, b) \\ = - \int_a^b \nabla u(s, \varphi_a) W(ds, \varphi_a) + \nabla u(a, \varphi_a) [W(b, \varphi_a) - W(a, \varphi_a)]. \end{aligned}$$

The estimate (2.4) (or (2.5)) implies

$$|\mu_1(a, b) + \mu_2(a, b)| \lesssim |b - a|^{2\tau}.$$

Since $2\tau > 1$, Lemma 2.3 yields $\mathcal{J}_0^t \mu_1 = -\mathcal{J}_0^t \mu_2$. This completes the proof after observing that the aforementioned identity is exactly the same as (3.27). \square

Remark 3.8. In the context of ordinary differential equation of the type

$$\frac{dX}{dt}(t, x) = b(t, X(t, x)),$$

with non-regular vector field b , existence and uniqueness and stability of regular Lagrangian flows were proved by R.J. DiPerna and P.-L. Lions ([13]) for Sobolev vector fields with bounded divergence. This result has been extended by L. Ambrosio ([2]) to BV coefficients with bounded divergence. In [9], it is shown that under slightly relaxed assumptions many of the ODE results of DiPerna-Lions theory can be recovered, from a priori estimates, similar to (3.20). The current paper proposes another extension of this theory, where the vector field is distribution (rough) in time (derivative of a Hölder continuous function) and smooth in space. It is also interesting to extend the results presented here for vector fields which are rougher in time (see e.g. [34] for the linear case) or which are both rough in time and in space.

4. FEYNMAN-KAC FORMULA - A PATHWISE APPROACH

In this section we shall study the stochastic parabolic equation with Hölder continuous noise in a Hölder random media (see equation (4.4) below). A feature of this problem is that for the noise we don't assume any Hölder continuity in time variable. To make up for lack of regularity in time, we assume some regularity on spatial variables. In this case, the method presented in this section works for each sample path of the noise.

Throughout the current section, T is a fixed positive time. To describe the noise, we introduce the following space. Let β be a fixed non-negative number. We say that f is in $C_\beta^{0,1+\alpha}([0, T] \times \mathbb{R}^d)$ if it belongs to $C([0, T], C_{\text{loc}}^{1+\alpha}(\mathbb{R}^d))$ and satisfies the following condition

$$(4.1) \quad [\nabla f]_{\beta, \alpha} := \sup_{\substack{t \in [0, T]; \\ x, y \in \mathbb{R}^d; x \neq y}} \frac{|\nabla f(t, x) - \nabla f(t, y)|}{|x - y|^\alpha (1 + |x|^\beta + |y|^\beta)} < \infty.$$

We notice that the condition (4.1) implies the growth conditions on ∇f and f . More precisely, one has

$$(4.2) \quad [\nabla f]_{\alpha+\beta, \infty} := \sup_{t \in [0, T], x \in \mathbb{R}^d} \frac{|\nabla f(t, x)|}{1 + |x|^{\alpha+\beta}} < \infty,$$

and

$$(4.3) \quad [f]_{\alpha+\beta+1, \infty} := \sup_{t \in [0, T], x \in \mathbb{R}^d} \frac{|f(t, x)|}{1 + |x|^{\alpha+\beta+1}} < \infty.$$

It is easy to see that $\|f\|_{C_\beta^{0,1+\alpha}} := [f]_{\alpha+\beta+1,\infty} + [\nabla f]_{\alpha+\beta,\infty} + [\nabla f]_{\beta,\alpha}$ forms a norm on $C_\beta^{0,1+\alpha}([0, T] \times \mathbb{R}^d)$. In the rest of this section, we denote

$$C_\beta^{0,1+\alpha^-} = \bigcap_{0 < \alpha' < \alpha} C_\beta^{0,1+\alpha'}([0, T] \times \mathbb{R}^d).$$

Similar to the classical Hölder spaces, the space of smooth functions is not dense in $C_\beta^{0,1+\alpha}([0, T] \times \mathbb{R}^d)$. However, we can still approximate a function in $C_\beta^{0,1+\alpha}([0, T] \times \mathbb{R}^d)$ by smooth functions with a little trade off in spatial regularity. More precisely, let η be function in $C_c^\infty(\mathbb{R}^{d+1})$ supported in $(-1, 1)^{d+1}$ and $\iint \eta(t, x) dt dx = 1$. For $\epsilon > 0$, we put $\eta_\epsilon(t, x) = \epsilon^{-d-1} \eta(\epsilon^{-1}(t, x))$. Let f be in $C_\beta^{0,1+\alpha}([0, T] \times \mathbb{R}^d)$, we define $f_\epsilon(t, x) = (f * \eta_\epsilon)(t, x)$. It is clear that f_ϵ belongs to $C_c^\infty(\mathbb{R}^{d+1})$. In addition, we have the following result.

Lemma 4.1. *For every $\alpha' < \alpha$, $[\nabla f_\epsilon - \nabla f]_{\beta,\infty}$ and $[\nabla f_\epsilon - \nabla f]_{\beta,\alpha'}$ converge to 0 as ϵ goes to 0.*

Proof. We have

$$\begin{aligned} |\nabla f_\epsilon(t, x) - \nabla f(t, x)| &\leq \iint |\nabla f(t, z) - \nabla f(t, x)| \eta_\epsilon(t, x - z) dt dz \\ &\leq [\nabla f]_{\beta,\alpha} \iint |x - z|^\alpha (1 + |x|^\beta + |z|^\beta) \eta_\epsilon(t, x - z) dt dz \\ &\lesssim [\nabla f]_{\beta,\alpha} \epsilon^\alpha (1 + |x|^\beta), \end{aligned}$$

which implied $[\nabla f_\epsilon - \nabla f]_{\beta,\infty} \rightarrow 0$. This also implies

$$|\nabla f_\epsilon(t, x) - \nabla f_\epsilon(t, y) - \nabla f(t, x) + \nabla f(t, y)| \lesssim [\nabla f]_{\beta,\alpha} \epsilon^\alpha (1 + |x|^\beta + |y|^\beta).$$

On the other hand

$$\begin{aligned} |\nabla f_\epsilon(t, x) - \nabla f_\epsilon(t, y)| &\leq \iint |\nabla f(t, x - z) - \nabla f(t, y - z)| \eta_\epsilon(t, z) dt dz \\ &\leq [\nabla f]_{\beta,\alpha} |x - y|^\alpha \iint (1 + |x - z|^\beta + |y - z|^\beta) \eta_\epsilon(t, z) dt dz \\ &\lesssim [\nabla f]_{\beta,\alpha} |x - y|^\alpha (1 + |x|^\beta + |y|^\beta), \end{aligned}$$

thus

$$|\nabla f_\epsilon(t, x) - \nabla f_\epsilon(t, y) - \nabla f(t, x) + \nabla f(t, y)| \lesssim [\nabla f]_{\beta,\alpha} |x - y|^\alpha (1 + |x|^\beta + |y|^\beta).$$

Interpolating these two bounds, we get

$$\begin{aligned} |\nabla f_\epsilon(t, x) - \nabla f_\epsilon(t, y) - \nabla f(t, x) + \nabla f(t, y)| \\ \lesssim [\nabla f]_{\beta,\alpha} \epsilon^{\alpha-\alpha'} |x - y|^{\alpha'} (1 + |x|^\beta + |y|^\beta) \end{aligned}$$

for every $\alpha' < \alpha$. This implies $[\nabla f_\epsilon - \nabla f]_{\beta,\alpha'} \rightarrow 0$. \square

In Section 5 we shall give conditions on the covariance of a Gaussian field $W(t, x)$ such that it is in $C_\beta^{0,1+\alpha}([0, T] \times \mathbb{R}^d)$.

Assume that W belongs to the space $C_\beta^{0,1+\alpha}([0, T] \times \mathbb{R}^d)$, throughout this section, we denote $W_n = W * \eta_{1/n}$. We consider the following parabolic equation with multiplicative noise:

$$(4.4) \quad \partial_t u + Lu + u \partial_t W = 0, \quad u(T, x) = u_T(x),$$

where the terminal function u_T is assumed to be measurable with polynomial growth and L is a second order differential operator of the form

$$(4.5) \quad L = \frac{1}{2} \sum_{i,j=1}^d a^{ij}(t, x) \partial_{x_i} \partial_{x_j} + \sum_{i=1}^d b^i(t, x) \partial_{x_i}.$$

Here the novelty is that we allow the coefficients $a^{ij}(t, x) = a^{ij}(t, x, W)$ and $b^i(t, x) = b^i(t, x, W)$ depend on W . Since we are going to solve the equation and to establish a Feynman-Kac type formula pointwise for W , we omit the explicit dependence of a^{ij} and b^i on W . Notice that with a time reversal $t \rightarrow T - t$, we can solve the stochastic parabolic equation with initial condition:

$$\partial_t u = Lu - u \partial_t W, \quad u(0, x) = u_0(x).$$

The stochastic differential equations with random coefficients have been studied in a large amount of papers. For example, it has been used in the modeling of the pressure in an oil reservoir with a log normal random permeability in [27] (see in particular the references therein). Recently, there have been great amount of research work on uncertainty quantization from the numerical computation community. Many different types of stochastic partial differential equations with random coefficients have been studied. Let us only mention the books [23], [50], and the references therein. Since the classical Feynman-Kac formula has already experienced many applications including the so-called Monte-Carlos particle approximation (see [11, 12]), we expect that the Feynman-Kac formula we obtained will be a significant addition to this literature in particular in the use of Monte-Carlo method for the computations.

We assume the following conditions on the operator L appearing in the equation (4.4).

- (L1) L is uniformly elliptic, that is there exist positive numbers λ and Λ such that

$$\lambda |\xi|^2 \leq \sum_{i,j=1}^d a^{ij}(t, x) \xi^i \xi^j \leq \Lambda |\xi|^2, \quad \forall \xi \in \mathbb{R}^d.$$

- (L2) For every t , the coefficients $a(t, \cdot)$ belong to $C_b^{2+\alpha}(\mathbb{R}^d)$ with bounded derivatives uniformly in t . That is

$$\sup_t \|a(t, \cdot)\|_{C_b^{2+\alpha}(\mathbb{R}^d)} \leq \Lambda.$$

- (L3) b is Lipschitz continuous and has linear growth, that is, there exists a positive constant $\kappa(b)$ such that

$$\begin{aligned} \sup_t |b^i(t, x)| &\leq \kappa(b)(1 + |x|), \quad \forall \xi \in \mathbb{R}^d, \\ \sup_t |b^i(t, y) - b^i(t, x)| &\leq \kappa(b)|y - x|, \quad \forall x, y \in \mathbb{R}^d. \end{aligned}$$

Under our conditions on W , it turns out that we can define the Feynman-Kac solution to equation (4.4), namely,

$$u(r, x) = \mathbb{E}^B \left[u_T(X_T^{r,x}) \exp \left\{ \int_r^T W(ds, X_s^{r,x}) \right\} \right],$$

where $\{X_s^{r,x}, s \geq r\}$ is the diffusion process generated by L starting from x at time r . More precisely, for every $r \leq t \leq T$ and $x \in \mathbb{R}^d$, let $X_t^{r,x}$ be the diffusion process given by the stochastic differential equation

$$(4.6) \quad dX_t^{i,r,x} = \sigma^{ij}(t, X_t^{r,x}) \delta B_t^j + b^i(t, X_t^{r,x}) dt, \quad X_r^{r,x} = x,$$

where σ is the square root matrix of a , namely, $a^{ij} = \sum_{k=1}^d \sigma^{ik} \sigma^{jk}$ and δB_t denotes the Itô differential. We will occasionally omit the index r, x and write X_s for $X_s^{r,x}$. Under conditions **(L1)**–**(L3)**, it is well-known that the diffusion process $X_t^{r,x}$ exists and has finite moments of all orders.

Equation (4.4) with W replaced by W_n is classic and one can obtain a smooth solution u_n (see for instance [36] where a more general situation is studied). The main result of the current section is to show that u_n converges to the Feynman-Kac solution u defined above. There are three main tasks to be accomplished:

- (i) One needs to define the nonlinear integration $\int W(ds, X_s)$. Since here W is only continuous in time, this integration is different from the Young integration considered in Section 2.
- (ii) One needs to show exponential integrability of $\int W(ds, X_s)$. In particular, the function u defined by Feynman-Kac formula is well-defined.
- (iii) One needs to show that the exponential functional of this integration is stable under approximations by smooth functions.

The outline of this section is as follows. In subsection 4.1, we define the nonlinear stochastic integration $\int W(ds, X_s)$ and show that it has finite moment of all orders. Exponential integrability is obtained if W has strictly sub-quadratic growth, namely, if α and β in (4.1)–(4.3) satisfy $\beta + \alpha < 1$. In subsection 4.2, we show that the Feynman-Kac solution is indeed a solution in certain sense. When W has more regularity in time such as in the case of Brownian sheets or fractional Brownian sheets, one can use this regularity to reduce the regularity requirement in space. This case is considered in subsection 4.3 when W satisfies the conditions in Section 2. Along the way, we will make use of some fundamental estimates for exponential moment of various norms of the diffusion X on finite intervals. These estimates are stated and proved in Appendix B.

In what follows, \mathbb{E} denotes the expectation with respect to a Brownian motion B , $\|\cdot\|_p$ denotes the L^p norm corresponding to \mathbb{E} .

4.1. Nonlinear Stochastic integral. Let $X_t^{r,x}$ satisfy (4.6) and let W be in $C_\beta^{0,1+\alpha}([0, T] \times \mathbb{R}^d)$. We shall define a new nonlinear integration $\int_r^T W(ds, X_s^{r,x})$. If W is differentiable in time, the natural definition for this type of integration is $\int_r^T \partial_t W(s, X_s^{r,x}) ds$. If W satisfies **(W)** then we can define it as in Section 2. However, in this section, Hölder continuity of W on t is not required. On the other hand, we shall use the crucial fact that $\{X_t^{r,x}, t \geq r\}$ is a semimartingale. We first give the following definition.

Definition 4.2. Let W_n be a sequence of smooth functions with compact support converging to W in $C_\beta^{0,1+\alpha}([0, T] \times \mathbb{R}^d)$. We define

$$(4.7) \quad \int_r^T W(ds, X_s^{r,x}) = \lim_n \int_r^T \partial_s W_n(s, X_s^{r,x}) ds$$

if the above limit exists in probability.

Of course, at the first glance, there is no reason for the limit in (4.7) to converge. We will show, however, that the above definition is well-defined, thanks to smoothing effect of the diffusion process $X_s^{r,x}$. Our first task is to find an appropriate representation for the integration $\int_r^T \partial_t W_n(s, X_s^{r,x}) ds$. To accomplish this, we consider the partial differential equation

$$(\partial_t + L_0)v_n(r, x) = -\partial_t W_n(r, x), \quad v(T, x) = -W_n(T, x),$$

where we recall that L is defined by (4.5) and

$$L_0 = L - b\nabla = \frac{1}{2} \sum_{i,j=1}^d a^{ij}(t, x) \partial_{x_i} \partial_{x_j}.$$

We could have chosen $L_0 = L$ but the above choice of L_0 will allow us to show exponential integrability later. Since W_n is a smooth function, the solution v_n is a strong solution which is at least twice differentiable in space and once differentiable in time. We then apply Itô formula to obtain

$$\begin{aligned} dv_n(s, X_s^{r,x}) &= (\partial_t + L)v_n(s, X_s^{r,x})ds + \sigma^{ij}(s, X_s^{r,x}) \partial_{x_i} v_n(s, X_s^{r,x}) \delta B_s^j \\ &= -\partial_t W_n(s, X_s^{r,x})ds - b(s, X_s^{r,x}) \cdot \nabla v_n(s, X_s^{r,x})ds \\ &\quad + \sigma^{ij}(s, X_s^{r,x}) \partial_{x_i} v_n(s, X_s^{r,x}) \delta B_s^j. \end{aligned}$$

Thus, it follows that

$$\begin{aligned} (4.8) \quad & \int_r^T \partial_t W_n(s, X_s^{r,x}) ds \\ &= W_n(T, X_T^{r,x}) + v_n(r, x) - \int_r^T b(s, X_s^{r,x}) \cdot \nabla v_n(s, X_s^{r,x}) ds \\ &\quad + \int_r^T \sigma^{ij}(s, X_s^{r,x}) \partial_{x_i} v_n(s, X_s^{r,x}) \delta B_s^j. \end{aligned}$$

Notice that the time derivative in W_n is transferred to the spatial derivative in v_n . The next task is to show that v_n and its derivative ∇v_n converge. This is accomplished by some estimates which are in the same spirit of the well-known Schauder estimates for parabolic equations in Hölder spaces. More precisely, we have

Lemma 4.3. *Suppose that W belongs to $C_{\text{loc}}^2(\mathbb{R}^{d+1})$ and satisfies*

$$[W]_{\beta_1, \infty} := \sup_{0 \leq t \leq T} \sup_{x \in \mathbb{R}^d} \frac{|\nabla W(t, x)|}{1 + |x|^{\beta_1}} < \infty$$

and

$$[W]_{\beta_2, \alpha} := \sup_{0 \leq t \leq T} \sup_{x \neq y} \frac{|\nabla W(t, x) - \nabla W(t, y)|}{|x - y|^\alpha (1 + |x|^{\beta_2} + |y|^{\beta_2})} < \infty$$

for some non-negative numbers β_1, β_2 . Let v be a strong solution with polynomial growth to the partial differential equation

$$(4.9) \quad (\partial_t + L_0)v = -\partial_t W, \quad v(T, x) = -W(T, x).$$

Let $t \mapsto \varphi_t$ be the diffusion process generated by L_0 , that is

$$(4.10) \quad \varphi_t^{r,x} = x + \int_r^t \sigma(s, \varphi_s^{r,x}) \delta B_s, \quad t \geq r.$$

Then v is uniquely defined and verifies

$$(4.11) \quad (v + W)(r, x) = -\mathbb{E} \int_r^T L_0 W(s, \varphi_s^{r,x}) ds.$$

In addition, the following estimates hold

$$(4.12) \quad \sup_{x \in \mathbb{R}^d} \frac{|(v + W)(r, x)|}{1 + |x|^{\beta_1}} \leq c(\beta_1, \lambda, \Lambda)[(T - r)^{1/2} + (T - r)][\nabla W]_{\beta_1, \infty},$$

$$(4.13) \quad \sup_{x \in \mathbb{R}^d} \frac{|\nabla(v + W)(r, x)|}{1 + |x|^{\beta_2}} \leq c(\alpha, \beta_2, \lambda, \Lambda)[(T - r)^{\alpha/2} + (T - r)^{\alpha/2+1/2}][\nabla W]_{\beta_2, \alpha},$$

and for every $\alpha' \in (0, \alpha)$,

$$(4.14) \quad \sup_{x \in \mathbb{R}^d} \frac{|\nabla(v + W)(r, x) - \nabla(v + W)(r, y)|}{(1 + |x|^{\beta_2} + |y|^{\beta_2})|x - y|^{\alpha'}} \leq c(\alpha', \alpha, \beta_2, \lambda, \Lambda)[(T - r)^{(\alpha - \alpha')/2} + (T - r)^{(\alpha - \alpha')/2+1/2}][\nabla W]_{\beta_2, \alpha}.$$

The proof of this result, even though lengthy, is straight forward and is provided in details in Appendix C.

Proposition 4.4. *Suppose that W belongs $C_\beta^{0,1+\alpha}([0, T] \times \mathbb{R}^d)$. Then there exists a C^1 -generalized solution v to the parabolic partial differential equation*

$$(4.15) \quad (\partial_t + L_0)v = -\partial_t W, \quad v(T, x) = -W(T, x),$$

such that for every $0 < \alpha' < \alpha$, the following estimates hold

$$(4.16) \quad [v + W]_{\alpha+\beta+1, \infty} \leq c(\alpha, \beta, \lambda, \Lambda)[\nabla W]_{\alpha+\beta, \infty},$$

$$(4.17) \quad [\nabla(v + W)]_{\beta, \infty} \leq c(\alpha, \beta, \lambda, \Lambda)[\nabla W]_{\beta, \alpha},$$

$$(4.18) \quad [\nabla(v + W)]_{\beta, \alpha'} \leq c(\alpha, \alpha', \beta, \lambda, \Lambda)[\nabla W]_{\beta, \alpha}.$$

As a consequence, v belongs to the space $C_\beta^{0,1+\alpha^-}([0, T] \times \mathbb{R}^d)$.

Proof. We recall that η is the bump function defined at the beginning of this section and $W_n = W * \eta_{1/n}$. Lemma 4.1 yields $[W_n - W]_{\beta, \infty}$ and $[W_n - W]_{\beta, \alpha}$ converge to 0 as $n \rightarrow \infty$. Thanks to linearity of the equation (4.15), $v_n - v_m$ is a strong solution to

$$(\partial_t + L_0)(v_n - v_m) = -\partial_t(W_n - W_m), \quad (v_n - v_m)(T, x) = (W_n - W_m)(T, x).$$

The results in Lemma 4.3 (with $\beta_1 = \beta_2 = \beta$) imply

$$[(v_n + W_n) - (v_m + W_m)]_{\beta, \infty} \lesssim [\nabla W_n - \nabla W_m]_{\beta, \infty},$$

$$[\nabla(v_n + W_n) - \nabla(v_m + W_m)]_{\beta, \infty} \lesssim [\nabla W_n - \nabla W_m]_{\beta, \alpha},$$

and for every $\alpha' \in (0, \alpha)$,

$$[\nabla(v_n + W_n) - \nabla(v_m + W_m)]_{\beta, \alpha'} \lesssim [\nabla W_n - \nabla W_m]_{\beta, \alpha}.$$

As a consequence, v_n is a Cauchy sequence in $C([0, T], C^1(K))$ for every compact set K in \mathbb{R}^d . Thus v_n converges to v in $C([0, T], C^1(K))$ for every compact set K . It is then straightforward to verify that v is a weak solution to (4.15). The estimates (4.16), (4.17) and (4.18) follow from a limiting argument. \square

Theorem 4.5. *Suppose that W belongs to $C_\beta^{0,1+\alpha}([0, T] \times \mathbb{R}^d)$. Let v be the $C_\beta^{0,1+\alpha'}$ -generalized solution to (4.15) constructed in Proposition 4.4. Then for every $t \in [r, T]$, the integration $\int_r^t W(ds, X_s^{r,x})$ is well-defined (in the sense of Definition 4.2). Moreover, it has moment of all positive orders and satisfies*

$$(4.19) \quad \int_r^t W(ds, X_s^{r,x}) = v(r, x) - v(t, X_t^{r,x}) \\ - \int_r^t b(s, X_s^{r,x}) \cdot \nabla v(s, X_s^{r,x}) ds + \int_r^t \sigma^{ij}(s, X_s^{r,x}) \partial_{x_i} v(s, X_s^{r,x}) \delta B_s^j.$$

Proof. We consider $W_n = W * \eta_{1/n}$ as in the proof of the previous proposition. It follows from Itô formula that (see (4.8))

$$\int_r^t \partial_t W_n(s, X_s^{r,x}) ds \\ = v_n(r, x) - v_n(t, X_t^{r,x}) - \int_r^t b(s, X_s^{r,x}) \cdot \nabla v_n(s, X_s^{r,x}) ds \\ + \int_r^t \sigma^{ij}(s, X_s^{r,x}) \partial_{x_i} v_n(s, X_s^{r,x}) \delta B_s^j.$$

Lemma 4.3 and Proposition 4.4 say that v_n (and its derivatives) has polynomial growth and converges in $C([0, T]; C_{\text{loc}}^{1+\alpha'}(\mathbb{R}^d))$ to v for every $\alpha' < \alpha$. Hence, the right hand side of the above formula is convergent in $L^p(\Omega)$ for every $p > 1$. Passing through the limit in n yields the equation (4.19). \square

Remark 4.6. To define $\int_r^t W(ds, X_s^{r,x})$, usually one needs some regularity of W on the temporal variable t . The equation (4.19) states that the requirement of the regularity on t can be transformed to the one on spatial variable x of another function v (defined by (4.9)). The use of v appears in many situations. If L_0 is replaced by L in the definition of v (e.g. equation (4.9)) and the terminal condition is replaced $v(0, x) = \delta(x - y)$ for any fixed y , then v corresponds to the transition density of the process X_s . This transition density is a fundamental concept in Markov processes and some other fields. It has also been used to simplify the proofs of a number of inequalities (see e.g. [14], [28]). The reason to use L_0 instead of L is that we don't need to assume condition on b to define v and that $\partial_i v$ will appear in (4.19) even we use L . The removal of temporal regularity also appears in other context. For example, to study the equation $dX_t = b(X_t) + dB_t$, the transformation $Y_t = X_t - B_t$ will satisfy $\dot{Y}_t = b(Y_t + B_t)$. The map $(t, x) \mapsto \int_0^t b(x + B_s) ds$, averaging along the trajectories of a Brownian motion, then has better regularity than that of b . In the field of stochastic differential equations, this phenomena has been observed by A. M. Davie in [10] and is recently studied in more depth in [5].

As a direct consequence, we obtain

Corollary 4.7. *Let W be in $C_\beta^{0,1+\alpha}([0, T] \times \mathbb{R}^d)$. Then for every $\alpha' < \alpha$, $p > 2$ and K compact subset of \mathbb{R}^d ,*

$$\| \int_r^T W(ds, X_s^{r,x}) - \int_r^T W(ds, X_s^{r,y}) - \int_r^T W_n(ds, X_s^{r,x}) + \int_r^T W_n(ds, X_s^{r,y}) \|_p \\ \leq C(\alpha, \alpha', \beta, \lambda, \Lambda, K, T, p) ([\nabla(W - W_n)]_{\beta, \infty} + [\nabla(W - W_n)]_{\beta, \alpha}) |x - y|^{\alpha'}$$

Proof. Fix $\alpha' < \alpha$, $p > 2$ and K compact subset of \mathbb{R}^d . We put $g(r, x) = \int_r^T W(ds, X_s^{r,x})$, $g_n(r, x) = \int_r^T W_n(ds, X_s^{r,x})$ and $h = v - v_n$. From (4.19),

$$\|g(r, x) - g(r, y) - g_n(r, x) + g_n(r, y)\|_p \leq I_1 + I_2 + I_3 + I_4,$$

where

$$\begin{aligned} I_1 &= |h(r, x) - h(r, y)| \\ I_2 &= \|h(T, X_T^{r,x}) - h(T, X_T^{r,y})\|_p \\ I_3 &= \int_r^T \|(b \cdot \nabla h)(s, X_s^{r,x}) - (b \cdot \nabla h)(s, X_s^{r,y})\|_p ds \\ I_4 &= \left\| \int_r^T (\sigma \nabla h)(s, X_s^{r,x}) - (\sigma \nabla h)(s, X_s^{r,y}) \cdot \delta B_s \right\|_p. \end{aligned}$$

Proposition 4.4 implies

$$|\nabla h(z)| \lesssim ([\nabla(W - W_n)]_{\beta, \infty} + [\nabla(W - W_n)]_{\beta, \alpha})(1 + |z|^\beta),$$

and

$$|\nabla h(x) - \nabla h(y)| \lesssim [\nabla(W - W_n)]_{\beta, \alpha}(1 + |x|^{\beta'} + |y|^{\beta'})|x - y|^{\alpha'}$$

where $\beta' = \beta + \alpha - \alpha'$. Therefore we can estimate

$$I_1 = \left| \int_0^1 \nabla h(\tau x + (1 - \tau)y) d\tau(x - y) \right| \lesssim \|W - W_n\| |x - y|,$$

$$I_2 = \left\| \int_0^1 \nabla h(\tau X_T^{r,x} + (1 - \tau)X_T^{r,y}) d\tau(X_T^{r,x} - X_T^{r,y}) \right\|_p \lesssim \|W - W_n\| |x - y|,$$

$$\begin{aligned} I_3 &\leq \int_r^T \| [b(s, X_s^{r,x}) - b(s, X_s^{r,y})] \nabla h(s, X_s^{r,x}) \|_p ds \\ &\quad + \int_r^T \| b(s, X_s^{r,y}) [\nabla h(s, X_s^{r,x}) - \nabla h(s, X_s^{r,y})] \|_p ds \\ &\lesssim \|W - W_n\| |x - y|^{\alpha'}, \end{aligned}$$

where we have used Hölder inequality. Similarly, we can estimate I_4 using Burkholder-Davis-Gundy inequality to get $I_4 \lesssim |x - y|^{\alpha'}$. From these bounds, the result follows. \square

Proposition 4.8. *Suppose W belongs to $C_\beta^{0,1+\alpha}([0, T] \times \mathbb{R}^d)$ with $\alpha + \beta < 1$. Then $\int_r^t W(ds, X_s^{r,x})$ is exponentially integrable uniformly over compact sets. More precisely, for every $\gamma > 0$, K compact subset of \mathbb{R}^d*

$$(4.20) \quad \sup_{x \in K} \mathbb{E} \exp \left\{ \gamma \int_r^t W(ds, X_s^{r,x}) \right\} < \infty$$

for all $\gamma > 0$.

Proof. From (4.19) it suffices to show for every $\gamma > 0$,

$$(4.21) \quad \sup_{x \in K} \mathbb{E} \exp \left\{ \gamma \int_r^t \sigma^{ij}(s, X_s^{r,x}) \partial_i v(s, X_s^{r,x}) dB_s^j \right\} < \infty,$$

$$(4.22) \quad \sup_{x \in K} \mathbb{E} \exp \left\{ \gamma \int_r^t b(s, X_s^{r,x}) \cdot \nabla v(s, X_s^{r,x}) ds \right\} < \infty,$$

$$(4.23) \quad \sup_{x \in K} \mathbb{E} \exp \{ \gamma |v(t, X_t^{r,x})| \} < \infty.$$

Let $0 < \theta < 2$. We claim that

$$(4.24) \quad \sup_{x \in K} \mathbb{E} \exp \left\{ \gamma \int_r^t |X_s^{r,x}|^\theta ds \right\} < \infty, \quad \forall \gamma > 0.$$

In fact, by Jensen inequality

$$\mathbb{E} \exp \left\{ \gamma \int_r^t |X_s^{r,x}|^\theta ds \right\} \leq (T-r)^{-1} \int_r^T \mathbb{E} e^{\gamma(T-r)|X_s^{r,x}|^\theta} ds.$$

The quality on the right hand side is finite thanks to (B.4).

For any martingale M_t with $\mathbb{E} e^{2\langle M \rangle_t} < \infty$ we have

$$\begin{aligned} \mathbb{E} e^{M_t} &= \mathbb{E} e^{M_t - \langle M \rangle_t} e^{\langle M \rangle_t} \\ &\leq \left\{ \mathbb{E} e^{2M_t - 2\langle M \rangle_t} \right\}^{1/2} \left\{ \mathbb{E} e^{2\langle M \rangle_t} \right\}^{1/2} = \left\{ \mathbb{E} e^{2\langle M \rangle_t} \right\}^{1/2}. \end{aligned}$$

Thus we have

$$\begin{aligned} \mathbb{E} \exp \left\{ \gamma \int_r^t \sigma^{ij}(s, X_s^{r,x}) \partial_i v(s, X_s^{r,x}) \delta B_s^j \right\} \\ \leq \left\{ \mathbb{E} \exp \left[2\gamma^2 \int_r^t (a^{ij} \partial_i v \partial_j v)(s, X_s^{r,x}) ds \right] \right\}^{1/2}. \end{aligned}$$

Taking into account the growth property of ∇v (see (4.17)) and a , we have

$$\sup_{x \in K} \mathbb{E} \exp \left[2\gamma^2 \int_r^t (a^{ij} \partial_i v \partial_j v)(s, X_s^{r,x}) ds \right] \lesssim \sup_{x \in K} \mathbb{E} \exp \left[c \int_r^t |X_s^{r,x}|^{2(\alpha+\beta)} ds \right],$$

which together with the previous claim shows (4.21) since $2(\alpha + \beta) < 2$. Similarly, since b has linear growth

$$\sup_{x \in K} \mathbb{E} \exp \left[\gamma \int_r^t b(s, X_s^{r,x}) \cdot \nabla v(s, X_s^{r,x}) ds \right] \lesssim \sup_{x \in K} \mathbb{E} \exp \left[c \int_r^t |X_s^{r,x}|^{1+\alpha+\beta} ds \right],$$

which shows (4.22) since $1 + \alpha + \beta < 2$.

Using the growth property of v , i.e. the estimate (4.16),

$$\mathbb{E} \exp [\gamma |v(t, X_t^{r,x})|] \lesssim \mathbb{E} \exp [c |X_t^{r,x}|^{1+\alpha+\beta}],$$

which shows (4.23). \square

Lemma 4.9. *Let W be in $C_\beta^{0,1+\alpha}([0, T] \times \mathbb{R}^d)$. Suppose $\alpha + \beta < 1$. For every $\gamma > 0$ and $r \in [0, T]$, we put $u(r, x) = \mathbb{E} \exp \left[\gamma \int_r^T W(ds, X_s^{r,x}) \right]$ and $u_n(r, x) = \mathbb{E} \exp \left[\gamma \int_r^T W_n(ds, X_s^{r,x}) \right]$. Then u_n converges to u in $C^{0,\alpha'}([0, T] \times K)$ for every $\alpha' < \alpha$ and K compact in \mathbb{R}^d .*

Proof. For a smooth function f , using fundamental theorem of calculus, we obtain

$$\begin{aligned} f(x) - f(a) - f(y) + f(b) &= \int_0^1 \int_0^1 f''(\xi) [\tau(x-y) + (1-\tau)(a-b)] d\eta d\tau (x-a) \\ &\quad + \int_0^1 f'(\theta) d\tau (x-a-y-b), \end{aligned}$$

where

$$\begin{aligned} \xi &= \tau\eta x + (1-\tau)\eta a + \tau(1-\eta)y + (1-\tau)(1-\eta)b, \\ \theta &= \tau y + (1-\tau)b. \end{aligned}$$

Thus, for every x, y in K , with $f(w) = \exp(\gamma w)$, we have

$$\begin{aligned} &u(r, x) - u_n(r, x) - u(r, y) + u_n(r, y) \\ &= \gamma^2 \mathbb{E} \int_0^1 \int_0^1 f(\xi) [\tau A(x, y) + (1-\tau)A_n(x, y)] d\eta d\tau B_n(x) \\ (4.25) \quad &+ \gamma \mathbb{E} \int_0^1 f(\theta) d\tau C_n(x, y), \end{aligned}$$

where

$$\begin{aligned} A(x, y) &= \int_r^T W(ds, X_s^{r,x}) - \int_r^T W(ds, X_s^{r,y}), \\ A_n(x, y) &= \int_r^T W_n(ds, X_s^{r,x}) - \int_r^T W_n(ds, X_s^{r,y}), \\ B_n(x) &= \int_r^T W(ds, X_s^{r,x}) - \int_r^T W_n(ds, X_s^{r,x}), \\ C_n(x, y) &= A(x, y) - A_n(x, y). \end{aligned}$$

The random variables ξ and η are linear combinations of these terms. From Proposition 4.8, we know that moments of $f(\xi)$ and $f(\theta)$ are bounded uniformly in x and τ, η . On the other hand, from Corollary 4.7, for every $\alpha' < \alpha$ and $p > 2$

$$\begin{aligned} \|A(x, y)\|_p &\lesssim |x - y|^{\alpha'}, \\ \sup_n \|A_n(x, y)\|_p &\lesssim |x - y|^{\alpha'}, \\ \lim_{n \rightarrow 0} \sup_{x \in K} \|B_n(x)\| &= 0, \end{aligned}$$

and

$$\|C_n(x, y)\|_p \lesssim ([\nabla(W - W_n)]_{\beta, \infty} + [\nabla(W - W_n)]_{\beta, \alpha}) |x - y|^{\alpha'}.$$

From (4.25), applying Hölder inequality and the above estimates for A, B, C we obtain

$$\begin{aligned} &|u(r, x) - u_n(r, x) - u(r, y) + u_n(r, y)| \\ &\lesssim [\sup_{x \in K} \|B_n(x)\|_p + [\nabla(W - W_n)]_{\beta, \infty} + [\nabla(W - W_n)]_{\beta, \alpha}] |y - x|^{\alpha'} \end{aligned}$$

for all x, y in K and $\alpha' < \alpha$. This completes the proof. \square

4.2. Feynman-Kac formula I. If W is a smooth function, then the classical Feynman-Kac formula asserts that

$$(4.26) \quad u(r, x) = \mathbb{E}^B \left[u_T(X_T^{r,x}) \exp \left(\int_r^T W(ds, X_s^{r,x}) \right) \right]$$

is the unique strong solution to (4.4). Indeed, suppose W is smooth and u is a strong solution to (4.4). Applying Itô formula to the process

$$t \mapsto u(t, X_t^{r,x}) \exp \left\{ \int_r^t \partial_t W(s, X_s^{r,x}) ds \right\}$$

we obtain

$$\begin{aligned} & \delta u(t, X_t^{r,x}) \exp \left\{ \int_r^t \partial_t W(s, X_s^{r,x}) ds \right\} \\ &= \exp \left\{ \int_r^t \partial_t W(s, X_s^{r,x}) ds \right\} (\partial_t + L + \partial_t W) u(t, X_t^{r,x}) dt \\ &+ \exp \left\{ \int_r^t \partial_t W(s, X_s^{r,x}) ds \right\} \sigma^{ij}(t, X_t^{r,x}) \partial_{x_i} u(t, X_t^{r,x}) \delta B_t^j \end{aligned}$$

Taking into account that $(\partial_t + L)u + \partial_t W u = 0$ and integrating over $[r, T]$, we have

$$\begin{aligned} & u_T(X_T^{r,x}) \exp \left\{ \int_r^T \partial_t W(s, X_s^{r,x}) ds \right\} - u(r, x) \\ &= \int_r^T \exp \left\{ \int_r^t \partial_t W(s, X_s^{r,x}) ds \right\} \sigma^{ij}(t, X_t^{r,x}) \partial_{x_i} u(t, X_t^{r,x}) \delta B_t^j \end{aligned}$$

Formula (4.26) is deduced by taking expectation on both sides.

Theorem 4.10. Assume W belongs to $C_\beta^{0,1+\alpha}([0, T] \times \mathbb{R}^d)$ with $\alpha + \beta < 1$. Let $W_n = W * \eta_{1/n}$. Let u_n be the solution to the parabolic equation

$$\partial_t u_n + L u_n + u_n \partial_t W_n = 0, \quad u_n(T, x) = u_T(x).$$

Let u be the function defined in (4.26). Then u_n converges to u in $C^{0,\alpha'}([0, T] \times K)$ for every $\alpha' < \alpha$ and K compact set in \mathbb{R}^d . As a consequence, u belongs to $C_{\text{loc}}^{0,\alpha'}([0, T] \times \mathbb{R}^d)$ for all $\alpha' < \alpha$.

Proof. We notice that

$$\begin{aligned} & u_n(r, x) - u(r, x) \\ &= \mathbb{E} \left\{ u_T(X_T^{r,x}) \left[\exp \left(\int_r^T W_n(ds, X_s^{r,x}) \right) - \exp \left(\int_r^T W(ds, X_s^{r,x}) \right) \right] \right\}. \end{aligned}$$

This together with Lemma 4.9 yield the theorem. \square

We notice that if f and g are locally Hölder continuous functions on \mathbb{R}^d with exponents α and γ respectively. Suppose that f has compact support and $\alpha + \gamma > 1$. Then we can define the Young integral

$$\int_{\mathbb{R}^d} f(x) g(d^j x) = \int_{\mathbb{R}^d} f(x) g(x_1, \dots, x_{j-1}, dx_j, x_{j+1}, \dots, x_n) d\hat{x}_j$$

where $\hat{x}_j = (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$.

We now show that if W is sufficiently regular in space, the Feynman-Kac solution u in (4.26) satisfies an equation derived from (4.4) by a change of variable. To better explain our procedure, let us first assume that W is smooth in space and time and u_T is also smooth. In such case, the equation (4.4) has unique smooth solution u such that

$$\partial_t u(t, x) + Lu(t, x) + u \partial_t W(t, x) = 0$$

for every $t \geq 0$ and $x \in \mathbb{R}^d$. We would like to obtain an equation of u such that the time derivative of W does not appear. To this end, we notice that

$$\partial_t u + u \partial_t W = e^{-W} \partial_t (u e^W).$$

Hence, multiply the equation with e^W and integrate in time, we obtain

$$(4.27) \quad u_t = e^{W_T - W_t} u_T + \int_t^T e^{W_s - W_t} L u_s ds.$$

In contrast with (4.4), the equation (4.27) does not contain the time derivative of W . One can also interpret (4.27) in weak sense. More precisely, the following result holds.

Theorem 4.11. *Assume W belongs to $C_\beta^{0,1+\alpha}([0, T] \times \mathbb{R}^d)$ with $\alpha + \beta < 1$. Let u be the function defined in (4.26). Then there is a sequence of smooth functions W_n with compact supports convergent to W in $C_\beta^{0,1+\alpha}([0, T] \times \mathbb{R}^d)$ and a sequence of u_n such that u_n converges to u uniformly over all compact sets. Moreover, for every test function $\varphi \in C_c^\infty(\mathbb{R}^d)$ the sequence*

$$\int_t^T \int_{\mathbb{R}^d} \partial_i [e^{W_n(s,x) - W_n(t,x)} \varphi(x)] a^{ij}(s, x) \partial_j u_n(s, x) dx ds$$

is convergent. If $\alpha > 1/2$, then we can identify the limit as

$$\int_t^T \int_{\mathbb{R}^d} \partial_i (e^{W(s,x) - W(t,x)} \varphi(x)) a^{ij}(s, x) u(s, d^j x) ds.$$

In such case, u verifies the equation

$$(4.28) \quad \begin{aligned} \int_{\mathbb{R}^d} u(t, x) \varphi(x) dx &= \int_{\mathbb{R}^d} e^{W(T,x) - W(t,x)} u_T(x) \varphi(x) dx \\ &+ \frac{1}{2} \int_t^T \int_{\mathbb{R}^d} \partial_i (e^{W(s,x) - W(t,x)} \varphi(x)) a^{ij}(s, x) u(s, d^j x) ds \\ &- \int_t^T \int_{\mathbb{R}^d} \partial_i \left(e^{W(s,x) - W(t,x)} \varphi(x) \left[b^i(s, x) - \frac{1}{2} \partial_j a^{ij}(s, x) \right] \right) u(s, x) dx ds. \end{aligned}$$

Proof. We recall that $W_n = W * \eta_{1/n}$ defined at the beginning of this section. Let u_n be the solution to the parabolic equation

$$\partial_t u_n + Lu_n + u_n \partial_t W_n = 0, \quad u_n(T, x) = -W_n(T, x).$$

Then it is easily verified that

$$\begin{aligned} \int_{\mathbb{R}^d} u_n(t, x) \varphi(x) dx &= \int_{\mathbb{R}^d} e^{W_n(T, x) - W_n(t, x)} u_n(T, x) \varphi(x) dx \\ &\quad + \frac{1}{2} \int_t^T \int_{\mathbb{R}^d} \partial_i [e^{W_n(s, x) - W_n(t, x)} \varphi(x)] a^{ij}(s, x) \partial_j u_n(s, x) dx ds \\ &\quad + \int_t^T \int_{\mathbb{R}^d} e^{W_n(s, x) - W_n(t, x)} \varphi(x) \left[b^i(s, x) - \frac{1}{2} \partial_j a^{ij}(s, x) \right] \partial_i u_n(s, x) dx ds. \end{aligned}$$

In other words,

$$\begin{aligned} \frac{1}{2} \int_t^T \int_{\mathbb{R}^d} \partial_i [e^{W_n(s, x) - W_n(t, x)} \varphi(x)] a^{ij}(s, x) \partial_j u_n(s, x) dx ds \\ = \int_{\mathbb{R}^d} u_n(t, x) \varphi(x) dx - \int_{\mathbb{R}^d} e^{W_n(T, x) - W_n(t, x)} u_n(T, x) \varphi(x) dx \\ + \int_t^T \int_{\mathbb{R}^d} \partial_i \left(e^{W_n(s, x) - W_n(t, x)} \varphi(x) \left[b^i(s, x) - \frac{1}{2} \partial_j a^{ij}(s, x) \right] \right) u_n(s, x) dx ds. \end{aligned}$$

Since φ has compact support, it is clear that all the terms on the right hand side are convergent. This implies that

$$\int_t^T \int_{\mathbb{R}^d} \partial_i [e^{W_n(s, x) - W_n(t, x)} \varphi(x)] a^{ij}(s, x) \partial_j u_n(s, x) dx ds$$

is convergent. In case $\alpha > 1/2$, by Theorem 4.10, this limit is convergent in the context of Young integrations. Hence, taking the limit yields (4.28). \square

Remark 4.12. (i) The use of Itô formula in subsection 4.1 is inspired from the work [19]. In that work, an Itô-Tanaka trick is applied to obtain some estimates to the commutator related to DiPerna-Lions' theory ([13]).

(ii) In the case W belongs to $C_{\text{loc}}^{0,2}([0, T] \times \mathbb{R}^d)$, the Itô-Tanaka formula (4.8) is negligible. In fact, using integration by part, one has

$$\int_r^T \partial_t W_n(s, X_s^{r, x}) ds = W_n(T, X_T^{r, x}) - W_n(r, x) - \int_r^T \nabla W_n(s, X_s^{r, x}) dX_s^{r, x}$$

where the last integral is in Stratonovich sense. By passing through the limit $n \rightarrow \infty$, we obtain

$$\int_r^T \partial_t W(s, X_s^{r, x}) ds = W(T, X_T^{r, x}) - W(r, x) - \int_r^T \nabla W(s, X_s^{r, x}) dX_s^{r, x}.$$

Assuming ∇W has linear growth in the spatial variable and $\nabla^2 W$ is globally bounded, one can also show exponential integrability

$$\mathbb{E}^B \exp \left[\int_r^T \partial_t W(s, X_s^{r, x}) ds \right] < \infty.$$

We consider u as in (4.26). Using the approximation as in the proof of Theorem 4.11, we can show that u verifies

$$\begin{aligned} \int_{\mathbb{R}^d} u(t, x) \varphi(x) dx &= \int_{\mathbb{R}^d} e^{W(T, x) - W(t, x)} u(T, x) \varphi(x) dx \\ &\quad + \int_t^T \int_{\mathbb{R}^d} L^* [e^{W(s, x) - W(t, x)} \varphi(x)] u(s, x) dx ds \end{aligned}$$

for all test functions φ in $C_c^\infty(\mathbb{R}^d)$, where L^* is the adjoint of L .

4.3. Feynman-Kac formula II. In previous subsections, to obtain the Feynman-Kac solution (4.26) (See Theorem 4.11) we assume that W is only continuous in time but satisfies (4.1)-(4.3) for $f = W$. This means that we suppose the first spatial derivatives of W exist and are Hölder continuous in order to compensate the lack of regularity in time. For many other stochastic processes (such as Brownian sheet or fractional Brownian sheets), W is Hölder continuous in time. In this case, we may use this time regularity to relax the regularity requirement on space variable. In this subsection we obtain a Feynman-Kac formula for the solution to (4.4) when W satisfies the conditions of the type given in Section 2. For example, we do not require W to possess first derivatives. More precisely, we assume $W : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ satisfies the following condition.

(FK) There are constants $\tau, \lambda \in (0, 1]$ and $\beta > 0$ such that

$$(4.29) \quad \tau + \frac{1}{2}\lambda > 1, \quad \beta + \lambda < 2$$

and such that the seminorm

$$\begin{aligned} &\|W\|_{\beta, \tau, \lambda} \\ (4.30) \quad &:= \sup_{\substack{0 \leq s < t \leq T \\ x, y \in \mathbb{R}^d; x \neq y}} \frac{|W(s, x) - W(t, x) - W(s, y) + W(t, y)|}{(1 + |x| + |y|)^\beta |t - s|^\tau |x - y|^\lambda} \\ &\quad + \sup_{\substack{0 \leq s < t \leq T \\ x \in \mathbb{R}^d}} \frac{|W(s, x) - W(t, x)|}{(1 + |x|)^{\beta + \lambda} |t - s|^\tau} + \sup_{\substack{0 \leq t \leq T \\ x, y \in \mathbb{R}^d; x \neq y}} \frac{|W(t, y) - W(t, x)|}{(1 + |x| + |y|)^\beta |x - y|^\lambda} \end{aligned}$$

is finite.

We continue to use the same notations introduced in previous subsections. For example, $X_t = X_t^{r, x}$ denotes the solution to the equation (4.6). The objectives of this subsection are to show that the expression defined by (4.26) is well-defined under the above condition **(FK)** and is the solution to (4.4).

From $\tau + \frac{1}{2}\lambda > 1$, it follows that there is a $\gamma \in (0, 1/2)$ such that $\tau + \gamma\lambda > 1$. Since X_t is Hölder continuous of exponent γ , from Proposition 2.4, we known that $\int_r^T W(ds, X_s^{r, x})$ is well-defined and

$$(4.31) \quad \left| \int_r^T W(ds, X_s) \right| \leq C(1 + \|X\|_\infty^\beta)(1 + \|X\|_\gamma^\lambda).$$

Since $\beta + \lambda < 2$, Lemma B.2 yields that

$$\mathbb{E} \exp \left\{ c \int_r^T W(ds, X_s) \right\} < \infty$$

for all $c \in \mathbb{R}$. Thus we have

Proposition 4.13. *Assume the conditions (L1)-(L3) are satisfied. Let (4.29)-(4.30) be satisfied. If there is an $\alpha_0 \in (0, 2)$ such that $|u_T(x)| \leq C_2 e^{C_1|x|^{\alpha_0}}$, then $u(r, x)$ defined by (4.26) is finite. Namely,*

$$(4.32) \quad u(r, x) = \mathbb{E}^B \left[u_T(X_T^{r,x}) \exp \left(\int_r^T W(ds, X_s^{r,x}) \right) \right]$$

is well-defined.

Now, let $W_n(t, x)$ be a sequence of functions in $C_0^\infty([0, T] \times \mathbb{R}^d)$ convergent to $W(t, x)$ under the norm $\|W\|_\infty + \|W\|_{\beta, \tau, \lambda}$. Denote $v_n(r, x) = \int_r^T W_n(ds, X_s^{r,x})$ and $v(r, x) = \int_r^T W(ds, X_s^{r,x})$ and $\tilde{v}_n(r, x) = v_n(r, x) - v(r, x)$. Thus, for any $0 \leq r < t \leq T$, we have

$$\begin{aligned} |\tilde{v}_n(t, x) - \tilde{v}_n(r, x)| &= \left| \int_t^T \tilde{W}_n(ds, X_s^{t,x}) - \int_r^T \tilde{W}_n(ds, X_s^{r,x}) \right| \\ &\leq \left| \int_r^t \tilde{W}_n(ds, X_s^{r,x}) \right| + \left| \int_t^T [\tilde{W}_n(ds, X_s^{t,x}) - \tilde{W}_n(ds, X_s^{r,x})] \right| \\ &=: I_1(r, t) + I_2(r, t). \end{aligned}$$

Applying the estimate in Proposition 2.4 to $\tilde{W}_n = W_n - W$, we obtain

$$\begin{aligned} \frac{I_1(r, t)}{(t-r)^\tau} &\leq \kappa \|\tilde{W}_n\|_{\tau, \lambda} (1 + \|X^{r,x}\|_\infty) [1 + \|X^{r,x}\|_\gamma (t-r)^{\lambda\gamma}] \\ &\leq C \|\tilde{W}_n\|_{\tau, \lambda} (1 + \|X^{r,x}\|_\infty) [1 + \|X^{r,x}\|_\gamma]. \end{aligned}$$

Lemma B.2 states that $\|X^{r,x}\|_\infty$ is bounded in L^p for any $p \geq 1$. Thus

$$(4.33) \quad \lim_{n \rightarrow \infty} \mathbb{E}^B \left\{ \sup_{0 \leq r < t \leq T} \left| \frac{I_1(r, t)}{(t-r)^\tau} \right|^p \right\} = 0$$

for any $p \geq 1$.

From Proposition 2.11 we have with $\tau + \theta\lambda\gamma > 1$,

$$\begin{aligned} I_2(r, t) &\leq C \|\tilde{W}_n\|_{\beta, \tau, \lambda} \|X^{t,x} - X^{r,x}\|_\infty^\lambda (T-t)^\tau \\ &\quad + C \|\tilde{W}_n\|_{\beta, \tau, \lambda} \|X^{t,x} - X^{r,x}\|_\infty^{\lambda(1-\theta)} (T-t)^{\tau+\theta\lambda\gamma} \\ (4.34) \quad &\leq C \|\tilde{W}_n\|_{\beta, \tau, \lambda} \|X^{t,x} - X^{r,x}\|_\infty^{\lambda(1-\theta)} [1 + \|X^{t,x} - X^{r,x}\|_\infty^{\lambda\theta}]. \end{aligned}$$

Notice that $X_s^{r,x} = X_s^{t, X_t^{r,x}}$. We have for any $p \geq 1$ and $\gamma' < 1$, by using the Markov property of the process $X_t^{r,x}$,

$$\begin{aligned} \mathbb{E} \sup_{0 \leq r < t \leq T} \frac{\|X^{t,x} - X^{r,x}\|_\infty^p}{(t-r)^{\gamma'p/2}} &= \mathbb{E} \sup_{0 \leq r < t \leq T} \frac{\|X^{t,x} - X^{t, X_t^{r,x}}\|_\infty^p}{(t-r)^{\gamma'p/2}} \\ &\leq C \mathbb{E} \sup_{0 \leq r < t \leq T} \frac{\|X_t^{r,x} - x\|_\infty^p}{(t-r)^{\gamma'p/2}} \leq C, \end{aligned}$$

where the last inequality follows from a similar argument as the proof of (B.5). Combining this with (4.34) implies

$$\mathbb{E} \sup_{0 \leq r < t \leq T} \left| \frac{I_2(r, t)}{(t-r)^{\gamma'\lambda(1-\theta)/2}} \right|^p \leq C.$$

Assume $\lambda/2 + \tau - 1 > 0$. For any $\tau' \in (0, \lambda/2 + \tau - 1)$ it is possible to find $\theta \in (0, 1)$ and $0 < \gamma < 1/2$ such that $\tau + \theta\lambda\gamma > 1$ and $\tau' < \gamma'\lambda(1 - \theta)/2$. We see that $v(\cdot, x)$ is Hölder continuous of exponent τ' and

$$\lim_{n \rightarrow \infty} \|v_n(\cdot, x) - v(\cdot, x)\|_{\tau'} = 0$$

uniformly in compact set K of \mathbb{R}^d . From (4.32) it is easy to see that

$$\lim_{n \rightarrow \infty} \|u_n(\cdot, x) - u(\cdot, x)\|_{\tau'} = 0$$

uniformly in compact set K of \mathbb{R}^d . Thus we have

Proposition 4.14. *Let W_n be a sequence of smooth functions such that W_n converges to W in the norm $\|W\|_\infty + \|W\|_{\beta, \tau, \lambda}$ and u_n is the solution to (4.4) and u is given by (4.32). Then for any $\tau' < \lambda/2 + \tau - 1$, $u(t, x)$ is Hölder continuous of exponent τ' in time variable t and on any compact set K of \mathbb{R}^d ,*

$$(4.35) \quad \lim_{n \rightarrow \infty} \|u_n(\cdot, x) - u(\cdot, x)\|_{\tau'} = 0$$

uniformly on $x \in K$.

If $\tau + \tau' > 1$, then for any $\varphi \in C_0^\infty(\mathbb{R}^d)$, we have

$$\int_0^t \int_{\mathbb{R}^d} u_n(s, x) (s, x) \varphi(x) \frac{\partial}{\partial s} W_n(s, x) ds dx$$

converges to the Young integral

$$\int_0^t \int_{\mathbb{R}^d} u(s, x) (s, x) \varphi(x) W(ds, x) dx.$$

It is obvious that the existence of $\tau' > 0$ such that $\tau + \tau' > 1$ and $\tau' < \lambda/2 + \tau - 1$ is equivalent to $\lambda + 4\tau > 4$. The above argument means that $u(t, x)$ is a weak solution to (4.4), in the sense of next theorem.

Theorem 4.15. *Assume the conditions (L1)-(L3) are satisfied and assume there is an $\alpha_0 \in (0, 2)$ such that $|u_T(x)| \leq C_2 e^{C_1|x|^{\alpha_0}}$. Let $\|W\|_{\beta, \tau, \lambda}$ defined by (4.30) be finite, where the Hölder exponents λ and τ and the growth exponent β satisfy*

$$(4.36) \quad \tau > 1/2, \quad \beta + \lambda < 2, \quad \lambda + 4\tau > 4.$$

Then u defined by (4.32) is a weak solution to (4.4) in the sense that u satisfies

$$(4.37) \quad \begin{aligned} \int_{\mathbb{R}^d} u(t, x) \varphi(x) dx &= \int_{\mathbb{R}^d} u_0(x) \varphi(x) dx + \int_0^t \int_{\mathbb{R}^d} u(s, x) L^* \varphi(x) dx ds \\ &+ \int_0^t \int_{\mathbb{R}^d} u(s, x) \varphi(x) W(ds, x) dx, \end{aligned}$$

where φ is any smooth function with compact support and where the last integral is a Young integral.

Remark 4.16. Equation (4.37) is the definition of the weak solution used in [32], [34] and [35].

5. ASYMPTOTIC GROWTH OF GAUSSIAN SAMPLE PATHS

In Sections 2, 3 and 4, we assume the pathwise Hölder continuity and pathwise growth conditions on W in order to define and to solve (partial) differential equations related to the nonlinear integral $\int W(ds, \varphi_s)$. For instance, the conditions (2.1), (4.1), (4.2), (4.3) are essential in various parts of the paper. In probability theory, it is usually hard to obtain properties for (almost) every sample path of a stochastic process from its average properties (from its probability law). In this section, we investigate these pathwise Hölder continuity and pathwise growth problems for a stochastic process. We shall focus on Gaussian random fields. However, our method works well for other processes provided they satisfy some suitable normal concentration inequalities (for instance, see the assumptions in Theorem 5.5).

Let W be a stochastic process on $[0, T] \times \mathbb{R}^d$. An application of our results yields the asymptotic growth of the quantity

$$I(\delta, R) = \sup_{t \in [0, T]} \sup_{|x|, |y| \leq R; |x-y| \leq \delta} \frac{|W(t, \square[x, y])|}{|x_1 - y_1|^{\lambda_1} \cdots |x_d - y_d|^{\lambda_d}}$$

as $R \rightarrow \infty$, where $W(t, \square[x, y])$ denotes the d -increment of $W(t, \cdot)$ over the rectangle $[x, y]$. More precise definition is given in Subsection 5.2. If R is fixed, the quality $I(\delta, R)$ is the objective in our previous work [30] via a multiparameter version of Garsia-Rodemich-Rumsey inequality.

Let us mention some historical remarks. (Pathwise) boundedness and continuity for stochastic processes have been studied thoroughly in literature. One of the central ideas is originated in an important early paper by Garsia, Rodemich and Rumsey (1970) [21]. This was developed further by Preston (1971, 1972) [44, 45], Dudley (1973) [15] and Fernique (1975) [16]. In these considerations, the parameter space \mathbf{T} is bounded and treated as a “single-dimension” object. For instance, the well-known Dudley bound

$$\mathbb{E} \sup_{s, t \in \mathbf{T}} |W(t) - W(s)| \lesssim \int_0^{d_W(s, t)} \sqrt{\log N(\mathbf{T}, d_W, \varepsilon)} d\varepsilon$$

yields modulus of continuity in terms of the entropy number $N(T, d, \varepsilon)$. This is extended to a more precise bound in terms of majorizing measure

$$\mathbb{E} \sup_{\substack{s, t \in \mathbf{T} \\ d_W(s, t) \leq \delta}} |W(t) - W(s)| \lesssim \sup_{t \in \mathbf{T}} \int_0^\delta \log^{1/2} \frac{1}{\mu(B_{d_W}(t, u))} du.$$

The majorizing-measure bound turns out to be necessary for processes which satisfy normal concentration inequalities. This result by M. Talagrand is the milestone in theory of Gaussian processes. We refer the readers to [41, Chapter 6] and references therein for details and more historical facts. See also Talagrand’s monograph [49] in which the role of majorizing measure is replaced by a variational quality called $\gamma_2(\mathbf{T}, d_W)$.

Estimates for the d -increment of W over a rectangle are quite different. Difficulties arise since $W(\square[s, t])$ does not behave nicely as increments. In particular, the corresponding entropic “metric”

$$(\mathbb{E} W(\square[s, t])^2)^{1/2}$$

does not satisfy the triangle inequality, but rather behaves like a volume metric. To elaborate this point, let us consider the two dimensional case:

$$\begin{aligned} W(\square[s, t]) &= W(s_2, t_2) - W(s_2, t_1) - W(s_1, t_2) + W(s_1, t_1) \\ &= \Delta_{[t_2, t_1]} W(s_2) - \Delta_{[t_2, t_1]} W(s_1) = \Delta_{[s_2, s_1]} \Delta_{[t_2, t_1]} W, \end{aligned}$$

where $\tilde{W}(s) := \Delta_{[t_2, t_1]} W(s) = W(s; t_2) - W(s; t_1)$. This product-like property is essential in our current approach (see for instance inequality (5.5) below). Alternatively, to obtain a sharp bound for the difference, one can repeatedly apply the Garsia-Rodemich-Rumsey inequality first $\Delta_{[s_2, s_1]} \tilde{W}$ and then to $\Delta_{[s_2, s_1]} \Delta_{[t_2, t_1]} W$. Indeed, for bounded parameter domains equipped with Lebesgue measure, this direction was developed by the authors in [30]. This idea, while might be feasible, seems to be more complicated in our current setting with general (unbounded) parameter domains equipped with a general measure.

In Subsection 5.1, we will prove a deterministic inequality, which is more precise than the multiparameter Garsia-Rodemich-Rumsey inequality obtained in [30]. We then apply it to obtain a majorizing-measure bound on the d -increments of stochastic processes in Subsection 5.2. Our formulations are benefited from the treatment in [41]. We however did not consider the necessary conditions for these bounds (i.e. lower bounds). Results in these two subsections are applicable to general stochastic processes.

Given a well-developed toolbox to treat the case when \mathbf{T} is bounded (or for example, R is fixed in $I(\delta, R)$), the asymptotic growth for $I(\delta, R)$ as $R \rightarrow \infty$ can be obtained using concentration inequalities for Gaussian processes. More precise results are given for fractional Brownian fields. This is done in Subsection 5.3.

5.1. A deterministic inequality. Throughout the current subsection, we put $\Psi(u) = \exp(u^2) - 1$. Suppose μ is a nonnegative measure on \mathbf{T} and X is a measurable function on \mathbf{T} . We define

$$[X]_{\Psi, (\mathbf{T}, \mu)} := \inf \left\{ \alpha > 0 : \int_{\mathbf{T}} \Psi \left(\frac{X(t)}{\alpha} \right) \mu(dt) \leq 1 \right\}.$$

When the parameter space \mathbf{T} and the measure μ are clear from the context, we often suppress them and write $[X]_{\Psi}$ instead. The following result, whose proof is given in [41, pg. 256-258], is an application of the Young inequality

$$ab \leq \int_0^a g(x) dx + \int_0^b g^{-1}(x) dx,$$

where g is a real-valued, continuous and strictly increasing function.

Lemma 5.1. *Let X and f be measurable functions on \mathbf{T} , μ be a nonnegative measure on \mathbf{T} . Assume that $[X]_{\Psi, (\mathbf{T}, \mu)}$ is finite and $0 < \int |f| d\mu < \infty$. Then*

$$\int_{\mathbf{T}} |X(t)f(t)| \mu(dt) \leq 3[X]_{\Psi, (\mathbf{T}, \mu)} \int_{\mathbf{T}} |f(t)| \log^{1/2} \left(1 + \frac{|f(t)|}{\int |f(s)| \mu(ds)} \right) \mu(dt).$$

We consider the case when \mathbf{T} has the form $\mathbf{T} = \mathbf{T}_1 \times \cdots \times \mathbf{T}_\ell$. A parameter t in \mathbf{T} has ℓ components, $t = (t_1, \dots, t_\ell)$. For each $i = 1, \dots, \ell$, the space \mathbf{T}_i is equipped with a metric d_i . We also denote $d^*(s, t) = d_1(s_1, t_1) \cdots d_\ell(s_\ell, t_\ell)$ for every s, t in \mathbf{T} . Let X be a function on \mathbf{T} . We define the ℓ -increment of X over a “rectangle” $[s, t]$

as

$$X(\square[s, t]) = \prod_{j=1}^{\ell} (I - V_{j,s})X(t).$$

In the above expression, I is the identity operator, $V_{j,s}$ is the substitution operator which substitutes the j -th component of a function on \mathbf{T} by s_j , more precisely,

$$V_{j,s}X(t) = X(t_1, \dots, t_{j-1}, s_j, t_{j+1}, \dots, t_{\ell}).$$

We refer to [30] for a more detailed description on this ℓ -increment.

For each i , $B^i(t_i, u)$ denotes the open ball with radius u in the metric space (\mathbf{T}_i, d_i) centered at t_i . For each t in \mathbf{T} , we denote $B(t, u) = B^1(t_1, u) \times \dots \times B^{\ell}(t_{\ell}, u)$. For each j , put $D_j = \sup_{s_j, t_j \in \mathbf{T}_j} d_j(s_j, t_j)$.

For each $i = 1, \dots, \ell$, let μ_i be a probability measure on \mathbf{T}_i . Let $k = (k_1, \dots, k_{\ell})$ be a multi-index in \mathbb{N}^{ℓ} . We define

$$\begin{aligned} \mu_k^i(t_i) &= \mu^i(B^i(t_i, D_i 2^{-k_i})), & \rho_{k_i}(t_i, \cdot) &= \frac{1}{\mu_k^i(t_i)} 1_{B^i(t_i, D_i 2^{-k_i})}(\cdot) \\ \mu_k(t) &= \prod_{i=1}^{\ell} \mu_k^i(t_i), & \rho_k(t, \cdot) &= \prod \rho_{k_i}(t_i, \cdot) \end{aligned}$$

and

$$(5.1) \quad M_k(t) = \int_{\mathbf{T}} \rho_k(t, u) X(u) \mu(du).$$

We use the notations $k+1 = (k_1+1, \dots, k_{\ell}+1)$, $k+1_j = (k_1, \dots, k_{j-1}, k_j+1, k_{j+1}, \dots, k_{\ell})$, $\hat{t}_i = (t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_{\ell})$ and $\hat{\mathbf{T}}_i = \mathbf{T}_1 \times \dots \times \mathbf{T}_{i-1} \times \mathbf{T}_{i+1} \times \dots \times \mathbf{T}_{\ell}$.

Theorem 5.2. *Let $\{X(t), t \in \mathbf{T}\}$ be a measurable function on \mathbf{T} . We put $\mu = \mu^1 \times \dots \times \mu^{\ell}$ and*

$$Z = \inf \left\{ \alpha > 0 : \iint_{\mathbf{T} \times \mathbf{T}} \Psi \left(\frac{X(\square[u, v])}{\alpha d^*(u, v)} \right) \mu(du) \mu(dv) \leq 1 \right\}.$$

Assume that D_j , $j = 1, \dots, d$, and Z are finite. Then, for every s, t in \mathbf{T} such that the integral

$$\int_0^{d_1(s_1, t_1)} du_1 \dots \int_0^{d_{\ell}(s_{\ell}, t_{\ell})} du_{\ell} \left(\log^{1/2} \frac{1}{\mu(B(s, u))} + \log^{1/2} \frac{1}{\mu(B(t, u))} \right)$$

is finite, $M_k(\square[s, t])$ converges to a limit, denoted by $X'(\square[s, t])$, as k_1, \dots, k_{ℓ} go to infinity. In addition, $X'(\square[s, t])$ satisfies

$$(5.2) \quad |X'(\square[s, t])| \leq C^{\ell} Z \int_0^{d_1(s_1, t_1)} du_1 \dots \int_0^{d_{\ell}(s_{\ell}, t_{\ell})} du_{\ell} \left(\log^{1/2} \frac{1}{\mu(B(s, u))} + \log^{1/2} \frac{1}{\mu(B(t, u))} \right).$$

Proof. Fix s, t in \mathbf{T} . We choose the multi-index n such that $D_j 2^{-n_j-1} \leq d_j(s_j, t_j) \leq D_j 2^{-n_j}$ for each $j = 1, \dots, \ell$. It suffices to show that the following series satisfies the bound in (5.2)

$$(5.3) \quad |M_n(\square[s, t])| + \sum_{k \geq n} |M_{k+1}(\square[s, t]) - M_k(\square[s, t])|.$$

We estimate the first term. Notice that we can write

$$M_n(\square[s, t]) = \iint_{\mathbf{T} \times \mathbf{T}} X(\square[u, v]) \rho_n(s, u) \rho_n(t, v) \mu(du) \mu(dv).$$

We consider the function $\{Y(u, v), (u, v) \in \mathbf{T} \times \mathbf{T}\}$ defined by

$$Y(u, v) = \begin{cases} \frac{X(\square[u, v])}{d^*(u, v)} & \text{when } d^*(u, v) \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

It is clear that

$$\begin{aligned} |M_n(\square[s, t])| &\leq \iint_{\mathbf{T} \times \mathbf{T}} |Y(u, v)| d^*(u, v) \rho_n(s, u) \rho_n(t, v) \mu(du) \mu(dv) \\ &\lesssim (D_1 2^{-n_1}) \cdots (D_\ell 2^{-n_\ell}) \iint_{\mathbf{T} \times \mathbf{T}} |Y(u, v)| \rho_n(s, u) \rho_n(t, v) \mu(du) \mu(dv), \end{aligned}$$

since the support of $\rho_n(s, \cdot) \rho_n(t, \cdot)$, $d^*(u, v) \lesssim (D_1 2^{-n_1}) \cdots (D_\ell 2^{-n_\ell})$. We now apply Lemma 5.1 to the functions Y and $\rho_n(s, \cdot) \otimes \rho_n(t, \cdot)$ on the product space $(\mathbf{T} \times \mathbf{T}, \mu \otimes \mu)$, observing that $Z = [Y]_\Psi$, $\iint \rho_n(s, \cdot) \rho_n(t, \cdot) = 1$ and $\rho_n(s, u) \rho_n(t, v) \leq (\mu_n(s, u) \mu_n(t, v))^{-1}$,

$$\begin{aligned} |M_n(\square[s, t])| &\lesssim Z (D_1 2^{-n_1}) \cdots (D_\ell 2^{-n_\ell}) \iint_{\mathbf{T} \times \mathbf{T}} \rho_n(s, u) \rho_n(t, v) \log^{1/2} (1 + \rho_n(s, u) \rho_n(t, v)) \mu(du) \mu(dv) \\ &\lesssim Z (D_1 2^{-n_1}) \cdots (D_\ell 2^{-n_\ell}) \log^{1/2} \left(1 + \frac{1}{\mu_n(s) \mu_n(t)} \right). \end{aligned}$$

Since $d^*(s, t) \asymp (D_1 2^{-n_1}) \cdots (D_\ell 2^{-n_\ell})$, this shows

$$(5.4) \quad |M_n(\square[s, t])| \lesssim Z \int_0^{d_1(s_1, t_1)} du_1 \cdots \int_0^{d_\ell(s_\ell, t_\ell)} du_\ell \log^{1/2} \left(\frac{1}{\mu(B(s, u))} + \frac{1}{\mu(B(t, u))} \right).$$

We now estimate each term in the sum appear in (5.3). We denote $\tau_0 k = k$ and recursively $\tau_j k = \tau_{j-1} k + 1_j$ for each $j = 1, \dots, \ell$. For example, $\tau_1 k = (k_1 + 1, k_2, \dots, k_\ell)$ and $\tau_\ell k = k + 1$. We then write

$$(5.5) \quad |M_{k+1}(\square[s, t]) - M_k(\square[s, t])| \leq \sum_{j=1}^{\ell} |M_{\tau_j k}(\square[s, t]) - M_{\tau_{j-1} k}(\square[s, t])|.$$

Note that the multi-indices $\tau_j k$ and $\tau_{j-1} k$ differs by exactly 1 unit at the j -th component. Without loss of generality, we consider the case

$$|M_{\tilde{k}}(\square[s, t]) - M_k(\square[s, t])|,$$

where $\tilde{k} = k + 1_\ell = (k_1, \dots, k_{\ell-1}, k_\ell + 1)$. We adopt the notations $w = (w', w_\ell)$ for every w in \mathbf{T} ,

$$\rho'_k(s', u') = \rho_{k_1}(s_1, u_1) \cdots \rho_{k_{\ell-1}}(s_{\ell-1}, u_{\ell-1})$$

and similarly for $\rho'_k(t', v')$. We then write

$$M_k(\square[s, t]) = M_k(\square^{\ell-1}[s', t'], s_\ell) - M_k(\square^{\ell-1}[s', t'], t_\ell)$$

and similarly for $M_{\tilde{k}}(\square[s, t])$. Thus

$$\begin{aligned}
 & |M_{\tilde{k}}(\square[s, t]) - M_k(\square[s, t])| \\
 & \leq |M_{k+1_\ell}(\square^{\ell-1}[s', t'], s_\ell) - M_k(\square^{\ell-1}[s', t'], s_\ell)| \\
 (5.6) \quad & + |M_{k+1_\ell}(\square^{\ell-1}[s', t'], t_\ell) - M_k(\square^{\ell-1}[s', t'], t_\ell)| \\
 & = I_1 + I_2.
 \end{aligned}$$

We only need to estimate I_1 since I_2 is analogous. We have

$$\begin{aligned}
 & M_k(\square^{\ell-1}[s', t'], s_\ell) \\
 & = \iint_{\mathbf{T} \times \mathbf{T}} X(\square^{\ell-1}[u', v'], v_\ell) \rho'_k(s', u') \rho'_k(t', v') \rho_{\ell+1}(s_\ell, u_\ell) \rho_\ell(s_\ell, v_\ell) \mu(du) \mu(dv)
 \end{aligned}$$

and similarly

$$\begin{aligned}
 & M_{\tilde{k}}(\square^{\ell-1}[s', t'], s_\ell) \\
 & = \iint_{\mathbf{T} \times \mathbf{T}} X(\square^{\ell-1}[u', v'], u_\ell) \rho'_k(s', u') \rho'_k(t', v') \rho_{\ell+1}(s_\ell, u_\ell) \rho_\ell(s_\ell, v_\ell) \mu(du) \mu(dv).
 \end{aligned}$$

Note how the dummy variables v_ℓ and u_ℓ have been switched between the two formulas. Hence

$$\begin{aligned}
 & |M_{\tilde{k}}(\square^{\ell-1}[s', t'], s_\ell) - M_k(\square^{\ell-1}[s', t'], s_\ell)| \\
 & \leq \iint_{\mathbf{T} \times \mathbf{T}} |X(\square^{\ell-1}[u', v'], u_\ell) \rho'_k(s', u') \rho'_k(t', v') \rho_{\ell+1}(s_\ell, u_\ell) \rho_\ell(s_\ell, v_\ell) \mu(du) \mu(dv)|
 \end{aligned}$$

Similarly to the term $M_n(\square[s, t])$ one can obtain

$$\begin{aligned}
 & |M_{\tilde{k}}(\square^{\ell-1}[s', t'], s_\ell) - M_k(\square^{\ell-1}[s', t'], s_\ell)| \\
 & \lesssim Z(D_1 2^{-k_1}) \cdots (D_\ell 2^{-k_\ell}) \log^{1/2} \left(1 + \frac{1}{\mu_{\tilde{k}}(s) \mu_k(s)} \right) \\
 & \lesssim Z(D_1 2^{-k_1}) \cdots (D_\ell 2^{-k_\ell}) \log^{1/2} \left(1 + \frac{1}{\mu_k(s)} \right).
 \end{aligned}$$

Therefore, combining altogether (5.5), (5.6) and the previous estimate, we get

$$|M_{k+1}(\square[s, t]) - M_k(\square[s, t])| \lesssim Z\ell(D_1 2^{-k_1}) \cdots (D_\ell 2^{-k_\ell}) \log^{1/2} \left(1 + \frac{1}{\mu_k(s)} \right),$$

and hence,

$$\begin{aligned}
 & \sum_{k \geq n} |M_{k+1}(\square[s, t]) - M_k(\square[s, t])| \\
 & \lesssim Z\ell(D_1 2^{-n_1}) \cdots (D_\ell 2^{-n_\ell}) \log^{1/2} \left(1 + \frac{1}{\mu_n(s)} \right) \\
 & \lesssim Z\ell \int_0^{d_1(s_1, t_1)} du_1 \cdots \int_0^{d_\ell(s_\ell, t_\ell)} du_\ell \log^{1/2} \left(\frac{1}{\mu(B(s, u))} + \frac{1}{\mu(B(t, u))} \right).
 \end{aligned}$$

Together with the bound for $M_n(\square[s, t])$ (inequality (5.4)) and (5.3), this completes the proof. \square

Remark 5.3. In Theorem 5.2, X' may not be defined as a function on \mathbf{T} , that is for each t in \mathbf{T} , there is no a priory reason for $X'(t)$ to be defined. However, in order to keep the representation compact, we have abused of notations and denote the limit as $X'(\square[s, t])$. This object is well-defined for every fixed s, t in \mathbf{T} .

5.2. Majorizing measure. We now suppose that X is a stochastic process with the probability space (Ω, \mathcal{F}, P) . We introduce the ℓ -fold volumetric

$$d^\ell(s, t) = (\mathbb{E}[X(\square[s, t])]^2)^{1/2}.$$

Assume that $\sup_{s, t \in \mathbf{T}} d^\ell(s, t)$ is finite. In addition, for each i , there exists a metric d_i on \mathbf{T}_i such that

$$d^\ell(s, t) \leq d_1(s_1, t_1) \cdots d_\ell(s_\ell, t_\ell).$$

This is not a restriction since such collection of metrics always exists. For instance, one can choose

$$d_1(s_1, t_1) = \sup_{\hat{s}_1, \hat{t}_1 \in \hat{\mathbf{T}}_1} d^\ell(s, t)$$

and recursively

$$d_k(s_k, t_k) = \sup_{\hat{s}_k, \hat{t}_k \in \hat{\mathbf{T}}_k} \frac{d^\ell(s, t)}{\prod_{i=1}^{k-1} d_i(s_i, t_i)}$$

with the convention $0/0 = 0$.

We denote Z as in Theorem 5.2, that is

$$(5.7) \quad Z = \inf \left\{ \alpha > 0 : \iint_{\mathbf{T} \times \mathbf{T}} \Psi \left(\frac{X(\square[u, v])}{\alpha d^*(u, v)} \right) \mu(du) \mu(dv) \leq 1 \right\}.$$

We assume that Z is finite almost surely.

Example 5.4. Suppose X is a centered Gaussian process. Then Z has exponential tail. More precisely $P(Z > u) \leq (e \log 2)^{1/2} u 2^{-u^2/2}$ for all $u > (2 + 1/\log 2)^{1/2}$. This comes from a standard argument by Chebyshev inequality and Hölder inequality, see [41, pg. 256-258] for details.

As an application of Theorem 5.2, we have

Theorem 5.5. *Let $\{X(t), t \in \mathbf{T}\}$ be a stochastic process such that Z , defined in (5.7), is finite a.s. Then X has a version X' such that for all $\omega \in \Omega$ and s, t in \mathbf{T}*

$$|X'(\omega, \square[s, t])| \leq C^\ell Z(\omega) \int_0^{d_1(s_1, t_1)} du_1 \cdots \int_0^{d_\ell(s_\ell, t_\ell)} du_\ell \left(\log^{1/2} \frac{1}{\mu(B(s, u))} + \log^{1/2} \frac{1}{\mu(B(t, u))} \right).$$

In particular, if $\mathbb{E}Z$ is finite, then

$$\mathbb{E} \sup_{d_i(s_i, t_i) \leq \delta_i, 1 \leq i \leq \ell} |X(\square[s, t])| \leq C^\ell (\mathbb{E}Z) \sup_{s \in \mathbf{T}} \int_0^{\delta_1} du_1 \cdots \int_0^{\delta_\ell} du_\ell \log^{1/2} \frac{1}{\mu(B(s, u))}.$$

Proof. First note that for every t, v in \mathbf{T}

$$(5.8) \quad (\mathbb{E}|X(t) - X(v)|^2)^{1/2} \leq \sum_{i=1}^{\ell} d_i(t_i, v_i).$$

We recall the notation $M_k(t)$ in (5.1). We have

$$\begin{aligned} \mathbb{E}|X(t) - M_k(t)| &\leq \int_{\mathbf{T}} \mathbb{E}|X(t) - X(v)|\rho_k(t, v)\mu(dv) \\ &\leq \int_{\mathbf{T}} \sum_{i=1}^{\ell} d_i(t_i, v_i)\rho_k(t, v)\mu(dv) \leq \sum_{i=1}^{\ell} D_i 2^{-k_i}. \end{aligned}$$

Together with Borel-Cantelli lemma, this shows for all $t \in \mathbf{T}$, $M_k(t)$ converges to $X(t)$ almost surely. On the other hand, Theorem 5.2 shows for all $s, t \in \mathbf{T}$, $M_k(\square[s, t])$ converges to a limit, denoted by $X'(\square[s, t])$. This implies $X(\square[s, t]) = X'(\square[s, t])$ almost surely. The result is now followed from Theorem 5.2. \square

5.3. Asymptotic growth. Let $W(t, x)$ be a continuous Gaussian process on $[0, T] \times \mathbb{R}^d$ with mean 0. As in the previous subsection, we define the d -fold volumetric

$$d(x, y) = \sup_{t \in [0, T]} (\mathbb{E}[W(t, \square[x, y])]^2)^{1/2}.$$

Without loss of generality, we assume there are metrics d_1, \dots, d_d on \mathbb{R} such that $d^*(x, y) = d_1(x_1, y_1) \dots d_d(x_d, y_d)$ satisfies $d(x, y) \leq d^*(x, y)$.

Let $\delta = (\delta_1, \dots, \delta_\ell)$ be in $(0, \infty)^\ell$. The notation $d^*(x, y) \leq \delta$ means $d_i(x_i, y_i) \leq \delta_i$ for all $i = 1, 2, \dots, \ell$. We denote $|x|^* = \max_{1 \leq i \leq d} d_i(0, x_i)$ for every $x \in \mathbb{R}^d$. We are interested in the asymptotic growth of the process

$$W^*(\delta, R) = \sup_{t \in [0, T]} \sup_{\substack{d^*(x, y) \leq \delta \\ |x|^*, |y|^* \leq R}} |W(t, \square[x, y])|$$

as R gets large and δ can range freely in a bounded neighborhood of 0. W^* also depends on T . However since T will always be fixed in our consideration, we suppress the dependence on T in our notations. We put

$$\begin{aligned} \mathcal{S}_R &= \{x \in \mathbb{R}^d : |x|^* \leq R\}, \\ m(\delta, R) &= \mathbb{E}W^*(\delta, R), \end{aligned}$$

and

$$\sigma(\delta, R) = \sup_{t \in [0, T]} \sup_{\substack{d^*(x, y) \leq \delta \\ x, y \in \mathcal{S}_R}} (E|W(t, \square[x, y])|^2)^{1/2}.$$

We first prove the following concentration inequality

Lemma 5.6. *For any $r > 0$,*

$$(5.9) \quad P\left(\frac{1}{\sigma(\delta, R)} |W^*(\delta, R) - m(\delta, R)| > r\right) \leq 2e^{-r^2/2}.$$

As a consequence,

$$(5.10) \quad \mathbb{E}\psi_\rho\left(\frac{|W^*(\delta, R) - m(\delta, R)|}{\sigma(\delta, R)}\right) \leq c_\rho < \infty$$

for every $\rho < 1/2$, where $\psi_\rho = \exp(\rho x^2)$.

Proof. It suffices to show (5.9). Let $\{X(u), u \in \mathbf{T}\}$ be a Gaussian process. Assume that \mathbf{T} is finite. The following concentration inequality is standard

$$(5.11) \quad P\left(\frac{1}{\sigma} \left| \sup_{u \in \mathbf{T}} |X(u)| - \mathbb{E} \left[\sup_{u \in \mathbf{T}} |X(u)| \right] \right| > r\right) \leq 2e^{-r^2/2},$$

for every $\sigma \geq \sup_{u \in \mathbf{T}} (EX^2(u))^{1/2}$. We refer to [38] or [41, Theorem 5.4.3] for a proof of (5.11). We now fix $(t_1, x_1), \dots, (t_m, x_m)$ in $[0, T] \times \mathbb{R}^d$ such that $d^*(x_j, x_k) \leq \delta$ and $|x_j|^*, |x_k|^* \leq R$ for all j, k . We denote $x_j \sqcup x_k$ the collection of points z in \mathbb{R}^d such that each component of z is the corresponding component of either x_j or x_k . We consider the centered Gaussian random process $X(t_i, x_j \sqcup x_k) := W(t_i, \square[x_j, x_k])$ indexed by the parameters $\{t_i\}_{1 \leq i \leq m}$ and $\{x_j \sqcup x_k\}_{1 \leq j, k \leq m}$. It is clear that

$$\mathbb{E}X^2(t_i, x_j \sqcup x_k) \leq \sigma^2(\delta, R).$$

Thus, the inequality (5.11) becomes

$$P\left(\frac{1}{\sigma(\delta, R)} \left| \sup_{i, j, k \leq m} |W(t_i, \square[x_j, x_k])| - \mathbb{E} \left[\sup_{i, j, k \leq m} |W(t_i, \square[x_j, x_k])| \right] \right| > r\right) \leq 2e^{-r^2/2}.$$

An approximation procedure yields (5.9). \square

Theorem 5.7. *With probability one,*

$$(5.12) \quad \sup_{\delta \in (0, 1]^\ell} \limsup_{R \rightarrow \infty} \frac{|W^*(\delta, R) - m(2\delta, 2R)|}{\sigma(2\delta, 2R) \sqrt{\log(\delta_1^{-1} \cdots \delta_\ell^{-1} \log R)}} \leq \sqrt{2}$$

Proof. We put $p(\delta, R) = \delta_1^{-1} \cdots \delta_\ell^{-1} (\log R)^2$ and consider the random variable

$$\Theta = \sup_{\delta \in (0, 1]^\ell, R \geq 1} \frac{1}{p(\delta, R)} \psi_\rho \left(\frac{1}{\sigma(2\delta, 2R)} |W^*(\delta, R) - m(2\delta, 2R)| \right).$$

For each multi-index $j = (j_1, \dots, j_\ell)$ in \mathbb{N}^ℓ , we denote $2^{-j} = (2^{-j_1}, \dots, 2^{-j_\ell})$. The notation $\delta \leq 2^{-j}$ means $\delta_i \leq 2^{-j_i}$ for all $i = 1, 2, \dots, \ell$. Then using the monotonicity of p , ψ_ρ , W^* and σ , and (5.10) we have

$$\begin{aligned} \mathbb{E}\Theta &\leq \sum_{k \in \mathbb{N}, j \in \mathbb{N}^\ell} \mathbb{E} \sup_{\substack{2^{-j-1} \leq \delta \leq 2^{-j} \\ 2^{k-1} \leq R \leq 2^k}} \frac{1}{p(\delta, R)} \psi_\rho \left(\frac{1}{\sigma(2\delta, 2R)} |W^*(\delta, R) - m(2\delta, 2R)| \right) \\ &\leq \sum_{k \in \mathbb{N}, j \in \mathbb{N}^\ell} \frac{1}{p(2^{-j}, 2^{k-1})} \mathbb{E} \psi_\rho \left(\frac{1}{\sigma(2^{-j}, 2^k)} |W^*(2^{-j}, 2^k) - m(2^{-j}, 2^k)| \right) \\ &\leq c_\rho \sum_{k \in \mathbb{N}, j \in \mathbb{N}^\ell} \frac{1}{p(2^{-j}, 2^{k-1})} < \infty. \end{aligned}$$

Hence, with probability one, Θ is finite and

$$\psi_\rho \left(\frac{1}{\sigma(2\delta, 2R)} |W^*(\delta, R) - m(2\delta, 2R)| \right) \leq \Theta p(2\delta, R), \quad \forall \delta > 0, \forall R \geq 1.$$

In particular,

$$\frac{|W^*(\delta, R) - m(2\delta, 2R)|}{\sigma(2\delta, 2R)} \leq \sqrt{\frac{\log[\Theta p(2\delta, R)]}{\rho}}, \quad \forall \delta > 0, \forall R \geq 1.$$

We then use the trivial estimate

$$\sqrt{\log(\Theta p)} \leq \sqrt{|\log \Theta|} + \sqrt{|\log p|}$$

to get

$$\frac{|W^*(\delta, R) - m(2\delta, 2R)|}{\sigma(2\delta, 2R)\sqrt{\log(\delta_1^{-1} \cdots \delta_\ell^{-1} \log R)}} \leq \sqrt{\frac{|\log \Theta|}{\rho |\log \log R|}} + \sqrt{\frac{\log(\delta_1^{-1} \cdots \delta_\ell^{-1} (\log R)^2)}{\rho \log(\delta_1^{-1} \cdots \delta_\ell^{-1} \log R)}},$$

for all $\delta > 0$ and $R \geq 1$. Since ρ can be chosen to be any constant less than $1/2$, we can choose a sequence ρ_n convergent to $1/2$. Since countable unions of events with probability zero still have probability zero, we can pass through the limit $n \rightarrow \infty$ to get, with probability one,

$$\frac{|W^*(\delta, R) - m(2\delta, 2R)|}{\sigma(2\delta, 2R)\sqrt{\log(\delta_1^{-1} \cdots \delta_\ell^{-1} \log R)}} \leq \sqrt{\frac{2|\log \Theta|}{|\log \log R|}} + \sqrt{\frac{2 \log(\delta_1^{-1} \cdots \delta_\ell^{-1} (\log R)^2)}{\log(\delta_1^{-1} \cdots \delta_\ell^{-1} \log R)}},$$

for all $\delta > 0$ and $R \geq 1$. Finally, let $R \rightarrow \infty$ to complete the proof. \square

In general, it is hard to say anything about the growth of $m(\delta, R)$ as R gets large. In what follows, we restrict ourselves to a particular (but still sufficiently large) class of Gaussian random fields. To be more precise, for each $i = 1, \dots, \ell$, let ϕ_i be a majorant for d_i , that is, ϕ_i is strictly increasing with $\phi_i(0) = 0$ and

$$(5.13) \quad d_i(x_i, y_i) \leq \phi_i(|y_i - x_i|).$$

Define

$$\tilde{\omega}_i(\delta_i) = \delta_i \log^{1/2} \frac{1}{\phi_i^{-1}(\delta_i)} + \int_0^{\phi_i^{-1}(\delta_i)} \frac{\phi_i(u)}{u \log^{1/2}(1/u)} du.$$

We will always presume $\tilde{\omega}_i$'s are finite wherever they appear.

Proposition 5.8. *Denote $\tilde{\delta}_i = \prod_{j \neq i} \delta_j$. Then we have*

$$(5.14) \quad m(\delta, R) \lesssim \delta_1 \cdots \delta_\ell \log^{1/2} \left(\prod_{i=1}^{\ell} 2\phi_i^{-1}(R) \right) + \sum_{i=1}^{\ell} \tilde{\delta}_i \tilde{\omega}_i(\delta_i)$$

where the implied constant is independent of R and δ .

Proof. We take for the majorizing measure $\mu_i = \lambda/(2\phi_i^{-1}(R))$, where λ is the Lebesgue measure. By (5.13), the ball $B^i(x_i, u_i)$ contains the interval $(x_i - \phi_i^{-1}(u_i), x_i + \phi_i^{-1}(u_i)) \cap \{z_i : d_i(z_i, 0) \leq R\}$, thus,

$$\mu_i(B^i(x_i, u_i)) \geq \frac{\phi_i^{-1}(u_i)}{2\phi_i^{-1}(R)}.$$

Hence, for all x in \mathcal{S}_R ,

$$\log \frac{1}{\mu(B(x, u))} \leq \log \left(\prod_{i=1}^d \frac{2\phi_i^{-1}(R)}{\phi_i^{-1}(u_i)} \right).$$

Therefore, for δ sufficiently small,

$$\begin{aligned} & \int_0^{\delta_1} du_1 \cdots \int_0^{\delta_\ell} du_\ell \log^{1/2} \frac{1}{\mu(B(x, u))} \\ & \leq \int_0^{\delta_1} du_1 \cdots \int_0^{\delta_\ell} du_\ell \log^{1/2} \left(\prod_{i=1}^d \frac{2\phi_i^{-1}(R)}{\phi_i^{-1}(u_i)} \right) \\ & \leq \delta_1 \cdots \delta_\ell \log^{1/2} \left(\prod_{i=1}^\ell 2\phi_i^{-1}(R) \right) + \int_0^{\delta_1} du_1 \cdots \int_0^{\delta_\ell} du_\ell \log^{1/2} \left(\prod_{i=1}^d \frac{1}{\phi_i^{-1}(u_i)} \right). \end{aligned}$$

The last integral in the above formula can be estimated as following

$$\begin{aligned} & \int_0^{\delta_1} du_1 \cdots \int_0^{\delta_\ell} du_\ell \log^{1/2} \left(\prod_{i=1}^d \frac{1}{\phi_i^{-1}(u_i)} \right) \\ & = \int_0^{\phi_1^{-1}(\delta_1)} d\phi_1(u_1) \cdots \int_0^{\phi_\ell^{-1}(\delta_\ell)} d\phi_\ell(u_\ell) \log^{1/2} \left(\prod_{i=1}^d \frac{1}{u_i} \right) \\ & \leq \sum_{i=1}^\ell \tilde{\delta}_i \int_0^{\phi_i^{-1}(\delta_i)} \log^{1/2}(1/u_i) d\phi_i(u_i). \end{aligned}$$

Using integration by parts, $\int_0^{\phi_i^{-1}(\delta_i)} \log^{1/2}(1/u_i) d\phi_i(u_i) \leq \tilde{\omega}(\delta_i)$, which completes the proof. \square

Example 5.9. Let $W = (W(x), x \in \mathbb{R}^d)$ be a fractional Brownian sheet with Hurst parameter $H = (H_1, \dots, H_d) \in (0, 1)^d$. In particular, the covariance of W is given by

$$\mathbb{E}W(x)W(y) = \prod_{i=1}^d R_{H_i}(x_i, y_i)$$

where

$$R_{H_i}(s, t) = \frac{1}{2}(|s|^{2H_i} + |t|^{2H_i} - |s - t|^{2H_i}).$$

We see that

$$(\mathbb{E}|W(\square[x, y])|^2)^{1/2} = \prod_{i=1}^d |x_i - y_i|^{H_i},$$

thus $\phi_i(\delta_i) = |\delta_i|^{H_i}$ and $\sigma(\delta, R) = \delta_1 \cdots \delta_d$. We put

$$m(\delta, R) = \mathbb{E} \sup |W(\square[x, y])|.$$

where the supremum is taken over the domain $\{x, y : |x_i|^{H_i}, |y_i|^{H_i} \leq R \text{ and } |x_i - y_i|^{H_i} \leq \delta_i \forall 1 \leq i \leq d\}$. Note that

$$\tilde{\omega}_i(\delta_i) \lesssim \delta_i \log^{1/2} \frac{1}{\phi_i^{-1}(\delta_i)}$$

The bound (5.14) yields

$$m(\delta, R) \lesssim \delta_1 \cdots \delta_d \sqrt{\log(R\delta_1^{-1} \cdots \delta_d^{-1})}.$$

Theorem 5.5 yields

$$(5.15) \quad \sup_{|x_i|^{H_i}, |y_i|^{H_i} \leq R; d^*(x, y) \leq \delta} |W(\square[x, y])| \lesssim \delta_1 \cdots \delta_d \sqrt{\log(R\delta_1^{-1} \cdots \delta_d^{-1})},$$

when R get large. This implies the inequality of the form (2.1) for W .

Remark 5.10. Fractional Brownian sheet belongs to a larger class of random fields called anisotropic random fields. That is,

$$(5.16) \quad (\mathbb{E}|W(y) - W(x)|^2)^{1/2} \asymp \sum_i |y_i - x_i|^{H_i}.$$

These random fields may have different behavior along different directions. In [42], the authors investigate the global moduli of continuity for anisotropic Gaussian random fields. As a result, they establish a sharp result for the global modulus of continuity for fractional Brownian sheets. The conditions considered in the current paper are somewhat more general. For instance, the estimate (5.8) implies the upper bound in the anisotropic condition (5.16). We believe our method (Theorems 5.2, 5.5) provides similar results as [42] though we do not report them here.

APPENDIX A. OTHER TYPES OF NONLINEAR STOCHASTIC INTEGRAL

The Itô integral is a fundamental concept in stochastic analysis. This integral can be defined under less condition than the Stratonovich one and has a completely different feature such as the famous Itô formula. From the modeling point of view, Itô type stochastic differential equations are more popular since all terms in the Itô equation $dx_t = b(x_t)dt + \sigma(x_t)\delta B_t$ (see also (4.6)) have clear meaning: $b(x_t)$ represents the mean rate of change and $\sigma(x_t)\delta B_t$ represents the fluctuation (it has zero mean contribution).

In this section, we will introduce nonlinear Itô-Skorohod integral. This integral is a probabilistic one and is defined for almost every sample path while nonlinear Young integral is defined for every sample path. The relation between these two integral is through the nonlinear symmetric (Stratonovich) integral.

A.1. Nonlinear Itô-Skorohod integral. Let $H \in (\frac{1}{2}, 1)$ and denote by $R_H(s, t) = \frac{1}{2}(s^{2H} + t^{2H} - |t - s|^{2H})$ the covariance function of a fractional Brownian motion of Hurst parameter H . Let $q(x, y)$ be a continuous and positive definite function, namely, for any $x_i \in \mathbb{R}^d$, $i = 1, 2, \dots, m$ and complex numbers $\xi_i, i = 1, 2, \dots, m$, not all 0, we have

$$\sum_{i,j=1}^m q(x_i, x_j) \bar{\xi}_i \xi_j \geq 0,$$

where $\bar{\xi}_i$ is the conjugate number of ξ_i . For every $s, t \geq 0$ and $x, y \in \mathbb{R}^d$, we denote

$$Q(s, t, x, y) = \frac{\partial^2 R_H}{\partial s \partial t}(s, t) q(x, y) = \alpha_H |s - t|^{2H-2} q(x, y),$$

where $\alpha_H = H(2H - 1)$. Let \mathcal{S} be the set of all smooth functions $f : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ such that $f(t, \cdot)$ has compact support for every $t \in [0, T]$. We introduce a scalar product on \mathcal{S} in the following way:

$$(A.1) \quad \langle \phi, \psi \rangle_{\mathcal{H}} = \int_{[0, T]^2 \times \mathbb{R}^{2d}} \phi(s, x) \psi(t, y) Q(s, t, x, y) dx dy ds dt.$$

We denote by \mathcal{H} the Hilbert space of the closure of \mathcal{S} with respect to this inner product. Let T be a bijective Hilbert-Schmidt operator on \mathcal{H} . Define the Banach space (in fact, it is a Hilbert space) Ω as the completion of \mathcal{H} with respect to the norm $\|x\|_{\Omega} := \sqrt{\langle Tx, Tx \rangle_{\mathcal{H}}}$. Then, it follows from the Bochner-Minlos theorem (see

[26], Theorem 3.1) that there is a probability measure P on (Ω, \mathcal{F}) such that $\langle h, \omega \rangle$ is a centered Gaussian random variable with covariance $\mathbb{E}[\langle h, \cdot \rangle \langle h', \cdot \rangle] = \langle h, h' \rangle_{\mathcal{H}}$, $\forall h, h' \in \Omega'$, where Ω' is the Banach space of all continuous linear functionals on Ω ; \mathcal{F} is the Borel σ -algebra generated by the open sets of Ω , and $\langle h, \omega \rangle$ the pairing between $h \in \Omega' \subset \mathcal{H}$ and Ω . We identify $\mathcal{H}' = \mathcal{H}$ so that the embeddings $\Omega' \subset \mathcal{H}' = \mathcal{H} \subset \Omega$ are continuous. We can define Gaussian random variable $\langle h, \omega \rangle$ for all $h \in \mathcal{H}$ by limiting argument.

First we give some specific elements in \mathcal{H} . For any $x \in \mathbb{R}^d$, we denote by δ_x the Dirac function on \mathbb{R}^d . Namely, δ_x is defined by $\int_{\mathbb{R}^d} \delta_x(y) f(y) dy = f(x)$ for any smooth function of compact support on \mathbb{R}^d .

Proposition A.1. *For any $s > 0$ and $x \in \mathbb{R}^d$, $I_{(0,s]} \delta_x$ is an element in \mathcal{H} and*

$$(A.2) \quad \langle I_{(0,s]} \delta_x, I_{(0,t]} \delta_y \rangle_{\mathcal{H}} = R_H(s, t) q(x, y)$$

and

$$(A.3) \quad \|I_{(0,s]} \delta_x - I_{(0,t]} \delta_y\|_{\mathcal{H}}^2 = s^{2H} q(x, x) + t^{2H} q(y, y) - 2R_H(s, t) q(x, y).$$

Proof. For every $\varepsilon > 0$ and $x \in \mathbb{R}^d$, we denote the elementary function

$$\delta_x^\varepsilon = (2\varepsilon)^{-d} I_{(x-\varepsilon, x+\varepsilon]}.$$

If ε tends to 0, the function δ_x^ε converges in \mathcal{H} to the generalized function δ_x . Indeed, fix (s, x) and (t, y) in $[0, T] \times \mathbb{R}^d$. For any positive numbers ε and ε' , we have

$$\langle I_{(0,s]} \delta_x^\varepsilon, I_{(0,t]} \delta_y^{\varepsilon'} \rangle_{\mathcal{H}} = R_H(s, t) (4\varepsilon\varepsilon')^{-d} \int_{y-\varepsilon'}^{y+\varepsilon'} \int_{x-\varepsilon}^{x+\varepsilon} q(x', y') dx' dy'.$$

Since $q(\cdot, \cdot)$ is continuous, the above right hand side converges to $q(x, y)$ as ε and ε' tend to 0. This shows easily that $I_{(0,s]} \delta_x^\varepsilon$ is a Cauchy sequence in \mathcal{H} when $\varepsilon \rightarrow 0$. The limit of $I_{(0,s]} \delta_x^\varepsilon$ in \mathcal{H} as $\varepsilon \rightarrow 0$ is $I_{(0,s]} \delta_x$. The equations (A.2) and (A.3) are immediate. \square

Since $I_{(0,s]} \delta_x \in \mathcal{H}$, we can define

$$(A.4) \quad W(s, x, \omega) = \langle I_{(0,s]} \delta_x, \omega \rangle, \quad \omega \in \Omega$$

Thus $\{W(s, x), t \geq 0, x \in \mathbb{R}^d\}$ is a multiparameter centered Gaussian process with the following covariance

$$\mathbb{E}[W(s, x)W(t, y)] = \langle I_{(0,s]} \delta_x, I_{(0,t]} \delta_y \rangle_{\mathcal{H}} = R_H(s, t) q(x, y).$$

We also denote

$$W(\phi) := \int_0^T \int_{\mathbb{R}^d} \phi(s, x) W(ds, x) dx := \langle \phi, \omega \rangle \quad \forall \phi \in \mathcal{H}.$$

We denote by \mathcal{P} the set of smooth and cylindrical random variables of the following form

$$(A.5) \quad F = f(W(\phi_1), \dots, W(\phi_n)),$$

$\phi_i \in \mathcal{H}$, $f \in C_p^\infty(\mathbb{R}^n)$ (f and all its partial derivatives have polynomial growth). D denotes the Malliavin derivative. That is, if F is of the form (A.5), then DF is the \mathcal{H} -valued random variable defined by

$$DF = \sum_{j=1}^n \frac{\partial f}{\partial x_j} (W(\phi_1), \dots, W(\phi_n)) \phi_j.$$

The operator D is closable from $L^2(\Omega)$ into $L^2(\Omega; \mathcal{H})$ and we define the Sobolev space $\mathbb{D}^{1,2}$ as the closure of \mathcal{P} under the norm

$$\|F\|_{1,2} = \sqrt{\mathbb{E}(F^2) + \mathbb{E}(\|DF\|_{\mathcal{H}}^2)}.$$

D can be extended uniquely to an operator from $\mathbb{D}^{1,2}$ into $L^2(\Omega; \mathcal{H})$. The *divergence operator* δ is the adjoint of the Malliavin derivative operator D . We say that a random variable u in $L^2(\Omega; \mathcal{H})$ belongs to the domain of the divergence operator, denoted by $\text{Dom } \delta$, if there is a constant $c_u \in (0, \infty)$ such that

$$|\mathbb{E}(\langle DF, u \rangle_{\mathcal{H}})| \leq c_u \|F\|_{L^2(\Omega)} \quad \forall F \in \mathbb{D}^{1,2}.$$

In this case $\delta(u)$ is defined by the duality relationship

$$(A.6) \quad \mathbb{E}(\delta(u)F) = \mathbb{E}(\langle DF, u \rangle_{\mathcal{H}}) \quad \forall F \in \mathbb{D}^{1,2}.$$

The following are two basic properties of the divergence operator δ .

(i) $\mathbb{D}^{1,2}(\mathcal{H}) \subset \text{Dom } \delta$ and for any $u \in \mathbb{D}^{1,2}(\mathcal{H})$

$$(A.7) \quad \mathbb{E}(\delta(u)^2) = \mathbb{E}(\|u\|_{\mathcal{H}}^2) + \mathbb{E}(\langle Du, (Du)^* \rangle_{\mathcal{H} \otimes \mathcal{H}}),$$

where $(Du)^*$ is the adjoint of Du in the Hilbert space $\mathcal{H} \otimes \mathcal{H}$.

(ii) For any F in $\mathbb{D}^{1,2}(\mathcal{H})$ and any u in the domain of δ such that Fu and $F\delta(u) - \langle DF, u \rangle_{\mathcal{H}}$ are square integrable, then Fu is in the domain of δ and

$$(A.8) \quad \delta(Fu) = F\delta(u) - \langle DF, u \rangle_{\mathcal{H}}.$$

The operator δ is also called the Skorokhod integral because in the case of Brownian motion, it coincides with the generalization of the Itô stochastic integral to anticipating integrands introduced by Skorokhod [47]. On the relation between δ and D , we have the identity

$$(A.9) \quad D\delta(u) = u + \delta(Du).$$

We refer to Nualart's book [43] for a detailed account of the Malliavin calculus with respect to a Gaussian process. Using the specific definition of our \mathcal{H} , we also denote $\delta(u) = \int_0^T \int_{\mathbb{R}^d} u(t, x) W(\delta t, x) dx$. In addition, we can write the identity (A.7) as

$$(A.10) \quad \begin{aligned} & \mathbb{E} \left[\int_0^T \int_{\mathbb{R}^d} u(t, x) W(\delta t, x) dx \right]^2 \\ &= \int_{[0, T]^2 \times \mathbb{R}^{2d}} \mathbb{E}[u(t, x)u(s, y)] Q(s, t, x, y) ds dt dx dy \\ &+ \int_{[0, T]^4 \times \mathbb{R}^{4d}} \mathbb{E}[D_{t_2, x_2} u(t_1, x_1) D_{s_2, y_2} u(s_1, y_1)] Q(t_1, s_2, x_2, y_1) Q(t_2, s_1, x_1, y_2) ds dt dx dy, \end{aligned}$$

where in the rest of the paper we shall use $ds = ds_1 \cdots ds_k$, $dx = dx_1 \cdots dx_m$ and so on, the k and m being clear in the context.

Let $\{W(t, x), t \geq 0, x \in \mathbb{R}^d\}$ be the Gaussian field introduced in Section A.1, whose mean is 0 and whose covariance is

$$\mathbb{E}(W(s, x)W(t, y)) = R_H(s, t)q(x, y).$$

Let $\varphi = \{\varphi_t, t \in [0, T]\}$ be a \mathbb{R}^d -valued stochastic process. Our aim in this section is to introduce and study the nonlinear stochastic integral $\int_0^T W(\delta t, \varphi_t)$.

This stochastic integral was studied earlier in order to establish the Feynman-Kac formula when φ_t is a Brownian motion, independent of W . The case $H > 1/2$ is discussed in [35] and the case $H < 1/2$ is discussed in [32]. When $\{W(t, x), t \geq 0\}$ is a semimartingale with respect to t (for fixed $x \in \mathbb{R}^d$), this type of stochastic integral has been studied extensively and generalized Itô formulas have been established. It has been applied to solve some stochastic partial differential equations. See for instance Kunita's book [37] and the references therein.

In this section, we will define the stochastic integral $\int W(\delta t, \varphi_t)$ based on the covariance structure of W . This method is closely tied to the nature of W as a Gaussian process. In particular, we introduce here two types of stochastic integrals, namely, the divergence type and symmetric type. We also study their properties and relation. The divergence type integral turns out to have zero mean, thus one can think of it as a generalization of Itô-Skorohod integral. The symmetric integral does not have vanishing mean and differs from the divergence type integral by a correction term, related to the Malliavin derivative of some random variable. One can also view the symmetric integral as a generalization of Stratonovich integral.

We shall define the (nonlinear) Itô-Skorohod (divergence) type integral $\int_0^T W(\delta t, \varphi_t)$ by the (linear) multi-parameter integral $\int_0^T \int_{\mathbb{R}^d} \delta(\varphi_t - y) W(\delta t, y) dy$. Here and in the remaining part of the paper, the symbol δ carries two meanings: the Itô-Skorohod integral and the Dirac delta function. Difference between the two meanings will be clear from the context.

Since $\delta(\varphi_t - y)$ is a distribution valued random process, to define its stochastic integral we need to approximate the Dirac delta function by smooth functions. Namely, we shall define $\int_0^T \int_{\mathbb{R}^d} \delta(\varphi_t - y) W(\delta t, y) dy$ as the limit $\lim_{\varepsilon \downarrow 0} \int_0^T \int_{\mathbb{R}^d} \eta_\varepsilon(\varphi_t - y) W(\delta t, y) dy$, where η_ε is an approximation of the Dirac delta function δ . To define such sequence η_ε , we denote by η the following bump function

$$\eta(x) = c_d \exp\{(|x|^2 - 1)^{-1}\} 1_{\{|x| < 1\}}, x \in \mathbb{R}^d,$$

where $|x|$ is the Euclidean distance in \mathbb{R}^d and c_d is the positive constant so that

$$\int_{\mathbb{R}^d} \eta(x) dx = 1.$$

The function η is smooth and compactly supported. Its corresponding mollifier is

$$(A.11) \quad \eta_\varepsilon(x) = \varepsilon^{-d} \eta\left(\frac{x}{\varepsilon}\right).$$

Here is our definition.

Definition A.2. Let $\varphi : [0, T] \times \Omega \rightarrow \mathbb{R}^d$ be a measurable stochastic process. If $I_\varepsilon = \int_0^T \int_{\mathbb{R}^d} \eta_\varepsilon(\varphi_t - y) W(\delta t, y) dy$ is well-defined and it has a limit in $L^2(\Omega, \mathcal{F}, P)$ as $\varepsilon \rightarrow 0$, then we define $\int_0^T W(\delta t, \varphi_t)$ as the aforementioned limit.

Next, we shall give condition to ensure the existence of the stochastic integral $\int_0^T W(\delta t, \varphi_t)$, namely, to ensure the existence of the limit of I_ε in $L^2(\Omega, \mathcal{F}, P)$. To express the conditions in a more concise way we introduce the following notations.

$$q_\varphi(x, y) = \alpha \int_0^T \int_0^T \mathbb{E} q(x + \varphi_s, y + \varphi_t) |s - t|^{2H-2} ds dt$$

and

$$q_{D\varphi}^*(x, y) = \alpha_H^2 \int_{[0, T]^4 \times \mathbb{R}^{2d}} \mathbb{E} D_{s_1, x'} q(x + \varphi_{s_2}, y') D_{t_2, y'} q(x', y + \varphi_{t_1}) \\ |s_1 - t_1|^{2H-2} |s_2 - t_2|^{2H-2} ds_1 ds_2 dt_1 dt_2 dx' dy'$$

whenever the integrals on the right hand side make sense. We make the following assumptions on the process φ_t .

- (A1) φ_t belongs to $\mathbb{D}^{1,2}$ for all t , and for almost every $\omega \in \Omega$, the sample path φ_t is continuous in $t \in [0, T]$.
- (A2) $|q|_\varphi$ is integrable on a neighborhood of $(0, 0)$, that is there exists an open set U in \mathbb{R}^{2d} containing $(0, 0)$ such that

$$\int_U \int_0^T \int_0^T \mathbb{E} |Q(s, t, x + \varphi_s, y + \varphi_t)| ds dt dx dy < \infty.$$

- (A3) $q_\varphi(x, y)$ is well-define in neighborhood of $(0, 0)$ and it is continuous at $(0, 0)$.
- (A4) There exists an open set U in \mathbb{R}^{2d} containing $(0, 0)$ such that

$$\int_U \int_{[0, T]^4 \times \mathbb{R}^{2d}} \mathbb{E} |D_{s_1, x'} q(x + \varphi_{s_2}, y') D_{t_2, y'} q(x', y + \varphi_{t_1})| \\ |s_1 - t_1|^{2H-2} |s_2 - t_2|^{2H-2} ds_1 ds_2 dt_1 dt_2 dx' dy' dx dy < \infty.$$

- (A5) $q_{D\varphi}^*(x, y)$ is well-defined in neighborhood of $(0, 0)$ and it is continuous at $(0, 0)$.

Theorem A.3. *We assume the conditions (A1)-(A5) are satisfied. Then $\int_0^T W(\delta t, \varphi_t)$ is well-defined and*

$$(A.12) \quad \mathbb{E} \left[\int_0^T W(\delta t, \varphi_t) \right]^2 = q_{D\varphi}^*(0, 0) + q_\varphi(0, 0) \\ = \int_{[0, T]^4} \int_{\mathbb{R}^{2d}} \mathbb{E} D_{s_1, x} Q(s_1, t_1, \varphi_{s_2}, y) D_{t_2, y} Q(s_2, t_2, x, \varphi_{t_1}) dx dy ds_1 ds_2 dt_1 dt_2 \\ + \int_0^T \int_0^T \mathbb{E} Q(s, t, \varphi_s, \varphi_t) ds dt.$$

Before proceeding to the proof, let us make the following remark which we will use several times in the future.

Remark A.4. Suppose that f and g are smooth functions, f has compact support, and φ is random variable in $\mathbb{D}^{1,2}$. Then the following integration by parts formula holds almost surely

$$(A.13) \quad \int_{\mathbb{R}^d} Df(x - \varphi) g(x) dx = - \int_{\mathbb{R}^d} f(x) Dg(x + \varphi) dx.$$

Indeed, the integration on the left hand side is

$$\int_{\mathbb{R}^d} \nabla f(x - \varphi) \cdot D\varphi g(x) dx.$$

Integrating by parts yields

$$- \int_{\mathbb{R}^d} f(x - \varphi) D\varphi \cdot \nabla g(x) dx.$$

With the change of the variable $x \mapsto x + \varphi$,

$$- \int_{\mathbb{R}^d} f(x) D\varphi \cdot \nabla g(x + \varphi) dx = - \int_{\mathbb{R}^d} f(x) Dg(x + \varphi) dx.$$

Proof of Theorem A.3. For any $\varepsilon > 0$, the \mathcal{H} -valued random variable $\eta_\varepsilon(\cdot - \varphi)$ belongs to $\mathbb{D}^{1,2}(\mathcal{H})$, hence belongs to $\text{Dom } \delta$. Thus, applying (A.7), for every positive numbers ε and ε' , we obtain

$$(A.14) \quad \mathbb{E}(\delta(\eta_\varepsilon(\cdot - \varphi))\delta(\eta_{\varepsilon'}(\cdot - \varphi))) = \mathbb{E}\langle \eta_\varepsilon(\cdot - \varphi), \eta_{\varepsilon'}(\cdot - \varphi) \rangle_{\mathcal{H}} \\ + \mathbb{E}\langle D\eta_\varepsilon(\cdot - \varphi), (D\eta_{\varepsilon'}(\cdot - \varphi))^* \rangle_{\mathcal{H} \otimes \mathcal{H}} =: E_1 + E_2.$$

Using a change of variable, we have

$$E_1 = \alpha_H \mathbb{E} \int_0^T \int_0^T \int_{\mathbb{R}^{2d}} \eta_\varepsilon(x - \varphi_s) \eta_{\varepsilon'}(y - \varphi_t) q(x, y) |t - s|^{2H-2} dx dy ds dt \\ = \alpha_H \mathbb{E} \int_0^T \int_0^T \int_{\mathbb{R}^{2d}} \eta_\varepsilon(x) \eta_{\varepsilon'}(y) q(x + \varphi_s, y + \varphi_t) |t - s|^{2H-2} dx dy ds dt \\ = \alpha_H \int_0^T \int_0^T \int_{\mathbb{R}^{2d}} \eta_\varepsilon(x) \eta_{\varepsilon'}(y) \mathbb{E} q(x + \varphi_s, y + \varphi_t) |t - s|^{2H-2} dx dy ds dt.$$

When ε and ε' tend to 0, using the conditions (A2), (A3), this quality converges to

$$\alpha_H \int_0^T \int_0^T \mathbb{E} q(\varphi_s, \varphi_t) |t - s|^{2H-2} ds dt = q_\varphi(0, 0).$$

Hence, when ε tends to zero, $\eta_\varepsilon(\cdot - \varphi)$ converges in $L^2(\Omega; \mathcal{H})$ to a \mathcal{H} -valued random variable, denoted by $\delta_\varphi = \delta(\varphi_t - y)$.

For the second expectation in (A.14), we use (A.10) to obtain

$$E_2 = \alpha_H^2 \mathbb{E} \int_{[0, T]^4 \times \mathbb{R}^{4d}} D_{s_1, x_1} \eta_\varepsilon(x_2 - \varphi_{s_2}) D_{t_2, y_2} \eta_{\varepsilon'}(y_1 - \varphi_{t_1}) q(x_1, y_1) q(x_2, y_2) \\ |s_1 - t_1|^{2H-2} |s_2 - t_2|^{2H-2} ds dt dx dy.$$

An application of (A.13) yields

$$E_2 = \alpha_H^2 \mathbb{E} \int_{[0, T]^4 \times \mathbb{R}^{4d}} \eta_\varepsilon(x_2) D_{s_1, x_1} q(x_2 + \varphi_{s_2}, y_2) \eta_{\varepsilon'}(y_1) D_{t_2, y_2} q(x_1, y_1 + \varphi_{t_1}) \\ |s_1 - t_1|^{2H-2} |s_2 - t_2|^{2H-2} ds dt dx dy.$$

When ε and ε' tend to 0, this converges to $q_{D\varphi}^*(0, 0)$ by using conditions (A4), (A5).

Therefore, $\delta(\eta_\varepsilon(\cdot - \varphi))$ is a Cauchy sequence in $L^2(\Omega)$. Since δ is a closed operator and $\eta_\varepsilon(\cdot - \varphi)$ converges to δ_φ , we obtain that δ_φ belongs to the domain of δ . As a consequence, $\delta(\eta_\varepsilon(\cdot - \varphi))$ converges to $\delta(\delta_\varphi)$ when ε tends to zero. Thus the integration $\int_0^T W(\delta t, \phi_t)$ is well-defined. The equation (A.12) is immediate. \square

Remark A.5. Under the hypothesis of the above theorem, the \mathcal{H} -valued random variable $\eta_\varepsilon(\cdot - \varphi_\cdot)$ converges in $L^2(\Omega; \mathcal{H})$ to $\delta_\varphi = \delta(\varphi_t - y)$ as ε tends to zero. Moreover, δ_φ also belongs to the domain of the divergence operator and the convergence also holds under the divergence δ . Hence, in this case, the stochastic integral in Definition A.2 can be viewed as $\delta(\delta_\varphi)$, the divergence of δ_φ .

A.2. Nonlinear symmetric stochastic integral. We introduce and study symmetric type stochastic integral by using appropriate approximation. This stochastic integral will be different than the Itô-Skorohod type integral introduced in the previous subsection.

Recall that $W = \{W(s, x, \omega), \omega \in \Omega\}$ is the Gaussian random field (indexed by (s, x)) defined in the previous subsection. Throughout this subsection, we assume that W is almost surely continuous with respect to $s \geq 0$ and $x \in \mathbb{R}^d$. We define the composition of the random field W and a \mathbb{R}^d -valued process $\varphi = \{\varphi_s, s \in [0, T]\}$ by

$$(A.15) \quad \begin{aligned} W(s, \varphi_s) : \Omega &\rightarrow \mathbb{R} \\ \omega &\mapsto W(s, \varphi_s(\omega), \omega). \end{aligned}$$

By convention, we will assume that all processes and functions vanish outside the interval $[0, T]$.

Definition A.6. The symmetric integral $\int_a^b W(d^{\text{sym}}s, \varphi_s)$ is defined as the limit as ε tends to zero of

$$(A.16) \quad (2\varepsilon)^{-1} \int_a^b (W(s + \varepsilon, \varphi_s) - W(s - \varepsilon, \varphi_s)) ds,$$

provided this limit exists in probability.

Example A.7. In the particular case, when $W(s, x) = B_s f(x)$, where f is a nice deterministic function and $\{B_s, s \geq 0\}$ is a Brownian motion, the symmetric integral defined above coincides with Stratonovich integral. That is $\int_0^T W(d^{\text{sym}}s, \varphi_s) = \int_0^T f(\varphi_s) d^\circ B_s$.

In the following proposition we will see that for a suitable class of \mathbb{R}^d -valued processes $\{\varphi_t\}$, the symmetric stochastic integral $\int_0^T W(d^{\text{sym}}s, \varphi_s)$ exists almost surely. This result is an extension of [1, Proposition 3].

Proposition A.8. *Let φ be a \mathbb{R}^d -valued process satisfying assumptions (A1)-(A5). In addition, suppose that φ satisfies*

$$(A.17) \quad \int_0^T \int_{|x| < 1} [\mathbb{E} q(x + \varphi_s, x + \varphi_s)]^{1/2} dx ds < \infty,$$

$$(A.18) \quad \int_0^T \int_{|x| < 1} \left[\mathbb{E} \left| \sum_{i,j=1}^d \langle D\varphi_s^i, D\varphi_s^j \rangle_{\mathcal{H}} \right| q(x + \varphi_s, x + \varphi_s) \right]^{1/2} dx ds < \infty$$

and the function

$$(A.19) \quad x \mapsto \int_0^T \int_0^T \int_{\mathbb{R}^d} |D_{t,y} q(x + \varphi_s, y)| |s - t|^{2H-2} dt ds dy$$

is a.s. well-defined and continuous on a neighborhood of 0. Assume also that the Gaussian field W has continuous sample path. Then the symmetric integral (A.16) exists and the following formula holds almost surely

$$(A.20) \quad \int_0^T W(d^{\text{sym}}s, \varphi_s) = \int_0^T W(\delta s, \varphi_s) + \alpha_H \int_0^T \int_0^T \int_{\mathbb{R}^{2d}} D_{t,y} q(\varphi_s, y) |s-t|^{2H-2} dt ds dy.$$

Proof. We shall show the convergence in L^2 of (A.16). For every positive ε , since W has continuous sample path, we can write

$$(A.21) \quad \begin{aligned} W(s+\varepsilon, \varphi_s) - W(s-\varepsilon, \varphi_s) &= \lim_{\varepsilon' \rightarrow 0} \int_{\mathbb{R}^d} [W(s+\varepsilon, x) - W(s-\varepsilon, x)] \eta_{\varepsilon'}(x - \varphi_s) dx \\ &= \lim_{\varepsilon' \rightarrow 0} \int_{\mathbb{R}^d} \delta(I_{[s-\varepsilon, s+\varepsilon]} \delta_x) \eta_{\varepsilon'}(x - \varphi_s) dx, \end{aligned}$$

almost surely, where we have used (A.4) in the last equality. We notice $\eta_{\varepsilon'}(x - \varphi_s)$ belongs to $\mathbb{D}^{1,2}$ for every s and x . Using (A.8), we see that the integrand on the right hand side of (A.21) can be written as

$$\delta(I_{[s-\varepsilon, s+\varepsilon]} \delta_x \eta_{\varepsilon'}(x - \varphi_s)) + \langle D\eta_{\varepsilon'}(x - \varphi_s), I_{[s-\varepsilon, s+\varepsilon]} \delta_x \rangle_{\mathcal{H}}.$$

Taking integration with respect to x and s , we obtain

$$(A.22) \quad \begin{aligned} &(2\varepsilon)^{-1} \int_0^T \int_{\mathbb{R}^d} [W(s+\varepsilon, x) - W(s-\varepsilon, x)] \eta_{\varepsilon'}(x - \varphi_s) dx ds \\ &= (2\varepsilon)^{-1} \int_0^T \int_{\mathbb{R}^d} \delta(I_{[s-\varepsilon, s+\varepsilon]} \delta_x \eta_{\varepsilon'}(x - \varphi_s)) dx ds \\ &\quad + (2\varepsilon)^{-1} \int_0^T \int_{\mathbb{R}^d} \langle D\eta_{\varepsilon'}(x - \varphi_s), I_{[s-\varepsilon, s+\varepsilon]} \delta_x \rangle_{\mathcal{H}} dx ds \\ &=: I_1 + I_2. \end{aligned}$$

The proof is now decomposed into several steps.

Step 1. Let us show that the integration with respect to $dx ds$ in I_1 can be interchanged with the divergence operator to obtain

$$I_1 = \delta \left((2\varepsilon)^{-1} \int_0^T \int_{\mathbb{R}^d} I_{[s-\varepsilon, s+\varepsilon]} \delta_x \eta_{\varepsilon'}(x - \varphi_s) dx ds \right).$$

In fact, one can view the integral in I_1 in Bochner sense, that is integration with L^2 -valued integrand. In this setting, we have

$$\int_0^T \int_{\mathbb{R}^d} \delta(u(s, x)) dx ds = \delta \left(\int_0^T \int_{\mathbb{R}^d} u(s, x) dx ds \right)$$

provided that

$$(A.23) \quad \int_0^T \int_{\mathbb{R}^d} \|u(s, x)\|_{\mathbb{D}^{1,2}} dx ds < \infty$$

and δ is a bounded operator from $\mathbb{D}^{1,2}$ to L^2 . The later fact is automatically guaranteed by (A.7). It remains to check that $u(s, x) = I_{[s-\varepsilon, s+\varepsilon]} \delta_x \eta_{\varepsilon'}(x - \varphi_s)$

satisfies (A.23).

$$\begin{aligned} \|u(s, x)\|_{\mathcal{H}}^2 &= \int_{s-\varepsilon}^{s+\varepsilon} \int_{s-\varepsilon}^{s+\varepsilon} \frac{\partial^2}{\partial s \partial t} R_H(s', t') ds' dt' q(x, x) \mathbb{E}[\eta_{\varepsilon'}^2(x - \varphi_s)] \\ &\leq R_H([0, T]^2) q(x, x) \mathbb{E}[\eta_{\varepsilon'}^2(x - \varphi_s)]. \end{aligned}$$

Thus by a change of variable, we obtain

$$\begin{aligned} \int_0^T \int_{\mathbb{R}^d} \|u(s, x)\|_{\mathcal{H}} dx ds &\leq R_H^{1/2}([0, T]^2) \int_0^T \int_{\mathbb{R}^d} (\mathbb{E} q(x, x) \eta_{\varepsilon'}^2(x - \varphi_s))^{1/2} dx ds \\ &= R_H^{1/2}([0, T]^2) \int_0^T \int_{\mathbb{R}^d} (\mathbb{E} q(x + \varphi_s, x + \varphi_s) \eta_{\varepsilon'}^2(x))^{1/2} dx ds \\ &\leq c(\varepsilon', T) \int_0^T \int_{|x| < 1} (\mathbb{E} q(x + \varphi_s, x + \varphi_s))^{1/2} dx ds. \end{aligned}$$

The last integral is finite thanks to the condition (A.17). Similarly

$$\begin{aligned} \|Du(s, x)\|_{\mathcal{H}}^2 &= \mathbb{E} \int_{s-\varepsilon}^{s+\varepsilon} \int_{s-\varepsilon}^{s+\varepsilon} \frac{\partial^2}{\partial s \partial t} R_H(s', t') ds' dt' q(x, x) \partial_i \eta_{\varepsilon'}(x - \varphi_s) \partial_j \eta_{\varepsilon'}(x - \varphi_s) \langle D\varphi_s^i, D\varphi_s^j \rangle_{\mathcal{H}} \\ &\leq R_H([0, T]^2) q(x, x) \mathbb{E} \left| \sum_{i,j} \partial_i \eta_{\varepsilon'}(x - \varphi_s) \partial_j \eta_{\varepsilon'}(x - \varphi_s) \langle D\varphi_s^i, D\varphi_s^j \rangle_{\mathcal{H}} \right|. \end{aligned}$$

Thus, by a change of variable and by using the condition (A.18), we obtain

$$\begin{aligned} \int_0^T \int_{\mathbb{R}^d} \|Du(s, x)\|_{\mathcal{H} \otimes \mathcal{H}} dx ds &\leq c(T) \int_0^T \int_{\mathbb{R}^d} [\mathbb{E} q(x, x) \left| \sum_{i,j} \partial_i \eta_{\varepsilon'}(x - \varphi_s) \partial_j \eta_{\varepsilon'}(x - \varphi_s) \langle D\varphi_s^i, D\varphi_s^j \rangle_{\mathcal{H}} \right|]^{1/2} dx ds \\ &\leq c(\varepsilon', T) \int_0^T \int_{|x| < 1} [\mathbb{E} q(x + \varphi_s, x + \varphi_s) \left| \sum_{i,j} \langle D\varphi_s^i, D\varphi_s^j \rangle_{\mathcal{H}} \right|]^{1/2} dx ds < \infty. \end{aligned}$$

Step 2. We show that

$$\begin{aligned} \delta \left((2\varepsilon)^{-1} \int_0^T \int_{\mathbb{R}^d} I_{(s-\varepsilon, s+\varepsilon]} \delta_x \eta_{\varepsilon'}(x - \varphi_s) dx ds \right) \\ = \delta \left((2\varepsilon)^{-1} \int_0^T I_{(s-\varepsilon, s+\varepsilon]} \eta_{\varepsilon'}(\cdot - \varphi_s) ds \right). \end{aligned}$$

It suffices to show for every smooth function ϕ with compact support

$$(A.24) \quad \phi = \int_{\mathbb{R}^d} \phi(y) \delta_y dy$$

is in \mathcal{H} , since with the choice $\phi = \eta_{\varepsilon'}$, (A.24) will yield the desired identity. Recall that \mathcal{S} is the space defined in Subsection A.1 and is dense in \mathcal{H} . Thus to show (A.24), we verify

$$\langle \phi, \psi \rangle_{\mathcal{H}} = \left\langle \int_{\mathbb{R}^d} \phi(y) \delta_y dy, \psi \right\rangle_{\mathcal{H}}$$

for every $\psi \in \mathcal{S}$. Indeed, we have

$$\begin{aligned} \left\langle \int_{\mathbb{R}^d} \phi(y) \delta_y dy, \psi \right\rangle_{\mathcal{H}} &= \int_{\mathbb{R}^d} \phi(y) \langle \delta_y, \psi \rangle_{\mathcal{H}} dy \\ &= \int_{\mathbb{R}^d} \int_0^T \int_0^T \int_{\mathbb{R}^d} \phi(y) \psi(x) Q(s, t, y, x) dx ds dt dy = \langle \phi, I_{(0,s]} \delta_x \rangle_{\mathcal{H}}, \end{aligned}$$

by Fubini's theorem.

Step 3. Combining the previous two steps, we obtain

$$I_1 = \delta \left((2\varepsilon)^{-1} \int_0^T I_{(s-\varepsilon, s+\varepsilon]} \eta_{\varepsilon'}(\cdot - \varphi_s) ds \right).$$

It is straightforward to check that when ε' and ε tend to zero, I_1 converges to $\int_0^T W(\delta s, \varphi_s)$ in L^2 .

Step 4. We now show the convergence of I_2 . A direct computation shows that

$$\begin{aligned} |I_2| &= (2\varepsilon)^{-1} \left| \int_0^T \int_{\mathbb{R}^d} \int_{-\varepsilon}^{\varepsilon} \int_0^T \int_{\mathbb{R}^d} D_{t,y} \eta_{\varepsilon'}(x - \varphi_s) q(x, y) |t - r - s|^{2H-2} dy dt dr dx ds \right| \\ &\leq d_H \int_0^T \int_{\mathbb{R}^d} \int_0^T \int_{\mathbb{R}^d} |D_{t,y} \eta_{\varepsilon'}(x - \varphi_s) q(x, y)| |t - s|^{2H-2} dy dt dx ds, \end{aligned}$$

where we have used the following inequality

$$(2\varepsilon)^{-1} \int_{-\varepsilon}^{\varepsilon} |t - r - s|^{2H-2} dr \leq d_H |t - s|^{2H-2}$$

for some constant d_H , independent $\varepsilon \in (0, 1)$ and $s, t \in \mathbb{R}$. By a change of variable $x - \varphi_s \rightarrow x$, we obtain

$$I_2 \leq d_H \int_0^T \int_0^T \int_{\mathbb{R}^{2d}} |\eta_{\varepsilon'}(x) D_{t,y} q(x + \varphi_s, y)| |t - s|^{2H-2} dy dx ds dt.$$

Hence, by the dominated convergence theorem, when ε' and ε tend to zero, I_2 goes to $\int_0^T \int_0^T \int_{\mathbb{R}^d} D_{t,y} q(x + \varphi_s, y) |t - s|^{2H-2} dy ds dt$. Therefore, passing through the limits in (A.22), we obtain (A.20) \square

If the limit in Definition A.6 exists for almost every sample path of W , then the symmetric integral can also be defined pathwise for a function $(W(t, x), t \geq 0, x \in \mathbb{R}^d)$. We also call such integral the symmetric integral and denoted by the same symbol $\int_0^T W(\mathfrak{d}^{\text{sym}} s, \varphi_s)$.

The following proposition establishes the relation between symmetric integral and nonlinear Young integral introduced in Section 2.

Proposition A.9. *Assume the hypothesis of Proposition 2.4. Then the symmetric integral exists and the following relation holds*

$$\int_0^T W(\mathfrak{d}^{\text{sym}} s, \varphi_s) = \int_0^T W(ds, \varphi_s).$$

Proof. Fix $\epsilon > 0$, we put

$$W_{\epsilon}(s, x) = (2\epsilon)^{-1} \int_{-\epsilon}^{\epsilon} W(s + \eta, x) d\eta.$$

We recall that $\int_0^T W(d^{\text{sym}}s, \varphi_s) = \lim_{\epsilon \rightarrow 0} \int_0^T \partial_t W_\epsilon(s, \varphi) ds$. We put

$$\begin{aligned}\mu_k(a, b) &= W_{\epsilon_k}(b, \varphi_a) - W_{\epsilon_k}(a, \varphi_a), \\ \mu(a, b) &= W(b, \varphi_a) - W(a, \varphi_a).\end{aligned}$$

Since W_ϵ is continuously differentiable in time, the integral $\int W_\epsilon(ds, \varphi_s)$ is understood in classical sense and is equal to $\int \partial_t W_\epsilon(s, \varphi_s) ds$. Hence, applying Proposition 2.10 we obtain, for any $\theta \in (0, 1)$ such that $\theta\tau + \lambda\gamma > 1$

$$\begin{aligned}& \left| \int_0^T W(ds, \varphi_s) - \int_0^T \partial_t W_\epsilon(s, \varphi_s) ds \right| \\ & \leq |W(T, \varphi_0) - W(0, \varphi_0) - W_\epsilon(T, \varphi_0) + W_\epsilon(0, \varphi_0)| \\ & \quad + c(\varphi)[W - W_\epsilon]_{\beta, \tau, \lambda} |b - a|^{\theta\tau + \lambda\gamma}.\end{aligned}$$

It remains to estimate the terms on the right side and show that they all converge to 0 when ϵ goes to 0. For the first term

$$\begin{aligned}& |W(T, \varphi_0) - W(0, \varphi_0) - W_\epsilon(T, \varphi_0) + W_\epsilon(0, \varphi_0)| \\ & \leq (2\epsilon)^{-1} \int_{-\epsilon}^\epsilon |W(T, \varphi_0) - W(0, \varphi_0) - W(T + \eta, \varphi_0) + W(\eta, \varphi_0)| d\eta \\ & \lesssim (2\epsilon)^{-1} \int_{-\epsilon}^\epsilon |\eta|^\tau d\eta \lesssim \epsilon^\tau.\end{aligned}$$

For the second term, we put $F = W - W_\epsilon$ and notice that

$$\begin{aligned}& |W_\epsilon(s, x) - W_\epsilon(s, y) - W_\epsilon(t, x) + W_\epsilon(t, y)| \\ & \leq (2\epsilon)^{-1} \int_{-\epsilon}^\epsilon |W(s + \eta, x) - W(s + \eta, y) - W(t + \eta, x) + W(t + \eta, y)| d\eta \\ & \leq [W](1 + |x|^\beta + |y|^\beta)(2\epsilon)^{-1} \int_{-\epsilon}^\epsilon |s - t|^\tau |x - y|^\lambda d\eta \\ & \leq [W](1 + |x|^\beta + |y|^\beta) |s - t|^\tau |x - y|^\lambda.\end{aligned}$$

Thus

$$|F(s, x) - F(s, y) - F(t, x) + F(t, y)| \leq 2[W](1 + |x|^\beta + |y|^\beta) |s - t|^\tau |x - y|^\lambda.$$

On the other hand,

$$\begin{aligned}& |W_\epsilon(s, x) - W_\epsilon(s, y) - W_\epsilon(t, x) + W_\epsilon(t, y)| \\ & \leq (2\epsilon)^{-1} \int_{-\epsilon}^\epsilon d\eta |W(s + \eta, x) - W(s, x) - W(s + \eta, y) + W(s, y)| \\ & \quad + |W(t, x) - W(t + \eta, x) - W(t, y) + W(t + \eta, y)| \\ & \leq 2[W](1 + |x|^\beta + |y|^\beta) |x - y|^\lambda (2\epsilon)^{-1} \int_{-\epsilon}^\epsilon |\eta|^\tau d\eta \\ & \leq 2(1 + \tau)^{-1} [W](1 + |x|^\beta + |y|^\beta) |x - y|^\lambda \epsilon^\tau.\end{aligned}$$

Hence, combining these two bounds, we get

$$\begin{aligned}& |F(s, x) - F(s, y) - F(t, x) + F(t, y)| \\ & \lesssim [W](1 + |x|^\beta + |y|^\beta) |s - t|^{\theta\tau} |x - y|^\lambda \epsilon^{\tau(1-\theta)}.\end{aligned}$$

Thus $[W - W_\epsilon]_{\beta, \theta\tau, \lambda} \lesssim \epsilon^{\tau(1-\theta)}$ which converges to 0 as $\epsilon \rightarrow 0$. \square

APPENDIX B. ESTIMATES FOR DIFFUSION PROCESS

In this section, we prove the exponential integrability of the Hölder norm and the supremum norm of a diffusion process which is needed in proving the existence of the Feynman-Kac solution in Section 4. The results obtained here are known in literature (see for instance [6], [18], [43]). However, it is difficult to find a single-source treatment that suits our purpose. Besides, our method is straightforward and unified. We present them here.

We recall that $X_t^{r,x}$ satisfies the equation (4.6). We denote

$$(B.1) \quad M_t^{r,x} = \sum_{j=1}^d \int_r^t \sigma^{ij}(s, X_s^{r,x}) \delta B_s^j.$$

Since σ is bounded (by condition (L1)), $(M_t^{r,x}; t \geq r)$ is a continuous L^2 martingale. In addition, we have the following properties.

Lemma B.1. *Let α be a number in $(0, 1/2)$. There exist some positive constants γ_0 and γ_α such that*

$$(B.2) \quad \mathbb{E} \exp \left\{ \gamma_0 \sup_{r \leq t \leq T} |M_t^{r,x}|^2 \right\} \leq C(T - r, \Lambda) < \infty$$

and

$$(B.3) \quad \mathbb{E} \exp \left\{ \gamma_\alpha \left(\sup_{r \leq s, t \leq T} \frac{|M_t^{r,x} - M_s^{r,x}|}{|t - s|^\alpha} \right)^2 \right\} \leq C(T - r, \Lambda, \alpha) < \infty.$$

Proof. (B.2) is well-known and is a direct application of Doob's maximal inequality and Burkholder-Davis-Gundy inequality. (B.3) is proved in [4, Lemma 2]. However, for readers' convenience, we present a proof of (B.3) in the following. We will omit the upper indices r, x . Applying the Garsia-Rodemich-Rumsey theorem (See [21] and [30], specifically [48, Theorem 2.1.3]) with $\Psi(x) = x^p$ and $p(x) = x^{\alpha+2/p}$, we have

$$|M_t - M_s| \leq 8(1 + \frac{2}{\alpha p}) 4^{1/p} |t - s|^\alpha \left\{ \int_r^T \int_r^T \left(\frac{|M_u - M_v|}{|u - v|^{\alpha+2/p}} \right)^p dudv \right\}^{1/p}.$$

Dividing both sides by $|t - s|^\alpha$ and taking the sup on $r \leq s < t \leq T$, we see that there is a constant $C = C(\alpha)$, independent of $p \geq 1$, such that

$$\mathbb{E} \left(\sup_{r \leq s < t \leq T} \frac{|M_t - M_s|}{|t - s|^\alpha} \right)^p \leq C^p \int_r^T \int_r^T \frac{\mathbb{E} |M_u - M_v|^p}{|u - v|^{\alpha p + 2}} dudv.$$

An application of the Burkholder-Davis-Gundy inequality gives

$$\|M_u - M_v\|_p \leq 2p^{1/2} \left\| \int_u^v a^{ii}(s, X_s^{r,x}) ds \right\|_{p/2}^{1/2} \leq 2\Lambda^{1/2} p^{1/2} (t - r)^{1/2}.$$

It follows that there is a constant C , which may be different than the above one, such that the p -moments of $\sup_{r \leq s < t \leq T} \frac{|M_t - M_s|}{|t - s|^\alpha}$ is at most $C^p p^{p/2} (T - r)^{(\frac{1}{2} - \alpha)p}$ for all $p > (\frac{1}{2} - \alpha)^{-1}$, which yields (B.3). \square

Lemma B.2. *Fix $\alpha \in (0, 1/2)$. There exist positive constants C_0 , γ_0 and γ_α such that*

$$(B.4) \quad \mathbb{E} \exp \left\{ \gamma_0 \sup_{r \leq t \leq T} |X_t^{r,x}|^2 \right\} \lesssim e^{C_0 |x|^2}$$

and

$$(B.5) \quad \mathbb{E} \exp \left\{ \gamma_\alpha \left(\sup_{r \leq s, t \leq T} \frac{|X_t^{r,x} - X_s^{r,x}|}{|t-s|^\alpha} \right)^2 \right\} \lesssim e^{C_0|x|^2}.$$

Proof. We denote $X_t^* = \sup_{r \leq s \leq t} |X_s^{r,x}|$ and $M_t^* = \sup_{r \leq s \leq t} |M_s^{r,x}|$. We first prove (B.4). Since b has linear growth (by (L3)), from equation (4.6), we see that

$$|X_t| \leq |M_t| + |x| + \kappa(b) \int_r^t |X_s| ds.$$

An application of Gronwall's inequality yields

$$|X_t| \leq |M_t| + |x| + \kappa(b) \int_r^t (|M_s| + |x|) e^{\kappa(b)(t-s)} ds.$$

Hence, for all $p \geq 0$, applying Jensen's inequality,

$$\begin{aligned} \exp\{pX_T^*\} &\leq \exp\{p(M_T^* + |x|)\} \exp\{p\kappa(b) \int_r^T (|M_s| + |x|) e^{\kappa(b)(t-s)} ds\} \\ &\leq \frac{\exp\{p(M_T^* + |x|)\}}{e^{\kappa(b)(T-r)} - 1} \int_r^T \exp\{p(e^{\kappa(b)(t-r)} - 1)(|M_s| + |x|) e^{\kappa(b)(T-s)}\} ds \\ &\lesssim \exp\{p(M_T^* + |x|)\} \int_r^T \exp\{Cp(|M_s| + |x|)\} ds \end{aligned}$$

for some constant C depending on $T-r$ and $\kappa(b)$. We then apply Cauchy-Schwartz inequality

$$\begin{aligned} \mathbb{E} e^{pX_T^*} &\lesssim \mathbb{E} e^{2p(M_T^* + |x|)} + \int_r^T \mathbb{E} e^{2Cp(|M_s| + |x|)} ds \\ &\lesssim \mathbb{E} e^{2Cp(M_T^* + |x|)}, \end{aligned}$$

where the constants (including the implied constant) are independent of p . Now we choose p according to the distribution $|N(0, a)|$ with a sufficient small, where $N(0, a)$ is a normal distribution independent of B . Using (B.2), the elementary estimate $\frac{1}{2}e^{\frac{a^2}{2}A^2} \leq \mathbb{E}^N e^{pA} \leq 2e^{\frac{a^2}{2}A^2}$ (with $A > 0$), and the previous estimate, we obtain (B.4).

From (4.6), we have

$$\frac{X_t - X_s}{(t-s)^\alpha} = \frac{\int_s^t b(u, X_u) du}{(t-s)^\alpha} + \frac{M_t - M_s}{(t-s)^\alpha} =: I_1 + I_2.$$

Since b has linear growth, $\sup_{r \leq s < t \leq T} |I_1| \leq c(\kappa(b), T, \alpha)(1 + X_T^*)$. (B.5) follows from (B.4) and (B.3). \square

APPENDIX C. SCHAUDER ESTIMATES

We present the proof of Lemma 4.3. The estimates (4.12)-(4.14) are similar to Schauder estimates in the classical theory of parabolic equations. Besides the results obtained in Appendix B, the method adopted here also makes use of Malliavin calculus. For this purpose, we need some preparations.

It is well-known (see e.g. [43]) that $X_t^{r,x}$ is differentiable (in Malliavin sense) with respect to the Brownian motion B_t . We denote the Malliavin derivative of

X with respect to B^j by $D^j X$. It is shown in [43, Theorem 2.2.1] that $DX = (D^1 X, \dots, D^d X)^T$ has finite moments of all orders and satisfies

$$dD_\tau X_t^{i,r,x} = \sigma_k^{ij}(t) D_\tau X_t^{k,r,x} \delta B_t^j + b_k^i(t) D_\tau X_t^{k,r,x} dt, \quad D_\tau^j X_\tau^{r,x} = \sigma^{ij}(\tau, X_\tau^{r,x})$$

for $t \geq \tau \geq r$, $D_\tau X_t^{r,x} = 0$ if $t < \tau \leq T$. In the above equation, we have used the notations

$$\sigma_k^{ij}(t) = \partial_{x_k} \sigma^{ij}(t, X_t^{r,x}), \quad b_k^i(t) = \partial_{x_k} b^i(t, X_t^{r,x}).$$

The matrix DX is understood as $[DX]^{ij} = D^j X^i$. Following the proof of [43, Theorem 2.2.1], one can show that the map $x \mapsto X_t^{r,x}$ is differentiable. We denote $Y(t; r, x) = \frac{\partial}{\partial x} X_t^{r,x}$, the Jacobian of $x \mapsto X_t^{r,x}$. The matrix Y is understood as $[Y]^{ij} = Y_j^i = \partial_j X^i$. It follows that the $d \times d$ -matrix valued process $t \mapsto Y(t; r, x)$ satisfies

$$(C.1) \quad \begin{aligned} dY_\bullet^i(t; r, x) &= \sigma_\bullet^{ij}(t) Y_\bullet^k(t; r, x) \delta B_t^j + b_\bullet^i(t) Y_\bullet^k(t; r, x) dt, \\ Y(r; r, x) &= I_{d \times d}. \end{aligned}$$

Let $Z(t)$ be the $d \times d$ matrix-valued process defined by

$$dZ_\bullet^\bullet(t) = -Z_\bullet^\bullet(t) \sigma_\bullet^{\theta l} \delta B_t^l - Z_\bullet^\bullet(t) [b_\bullet^\theta(t) - \sigma_\bullet^{\alpha l}(t) \sigma_\alpha^{\theta l}(t)] dt, \quad Z(r; r, x) = I_{d \times d}.$$

By means of Itô's formula, we have

$$\begin{aligned} d(Z_\bullet^k Y_j^i) &= -Y_j^i Z_\bullet^k \sigma_\bullet^{\theta l} \delta B_t^l - Y_j^i Z_\bullet^k b_\bullet^\theta dt + Z_\bullet^k Y_j^i \sigma_\bullet^{\alpha l} \sigma_\alpha^{\theta l} dt \\ &\quad + Z_\bullet^k Y_j^\theta \sigma_\bullet^{il} \delta B_t^l + Z_\bullet^k Y_j^\theta b_\bullet^i dt - Z_\bullet^k Y_j^\alpha \sigma_\bullet^{ij} \sigma_\alpha^{\theta l} dt = 0 \end{aligned}$$

and similarly for $Y_t Z_t$. Thus we obtain $Y_t Z_t = Y_t Z_t = I$. As a consequence, for every $t \geq r$, the matrix $Y(t; r, x)$ is invertible and its inverse is $Z(t; r, x)$. It is a standard fact that Y and Z have finite moments of all orders. More precisely, one has

$$(C.2) \quad \sup_{t \in [r, T], x \in \mathbb{R}^d} \mathbb{E} [|Y(t; r, x)|^p + |Y^{-1}(t; r, x)|^p] \leq c(p, T).$$

Since the coefficients of L are twice differentiable with bounded derivatives, DY exists and has finite moment of all orders and

$$(C.3) \quad \sup_{t \in [r, T], x \in \mathbb{R}^d} \mathbb{E} \sup_{\tau \in [r, T]} [|D_\tau Y(t; r, x)|^p + |D_\tau Y^{-1}(t; r, x)|^p] \leq c(p, T).$$

Moreover, it is well-known that the following representation holds (see, for instance [43, pg. 126])

$$(C.4) \quad D_\tau X_t^{r,x} = Y(t; r, x) Z(\tau; r, x) \sigma(\tau, X_\tau^{r,x}), \quad \forall \tau \in [r, t].$$

As a consequence, if f is a smooth function, we have

$$(C.5) \quad D_\tau f(s, X_s^{r,x})^T = \nabla f(s, X_s^{r,x})^T Y(s; r, x) Y^{-1}(\tau; r, x) \sigma(\tau, X_\tau^{r,x}).$$

(where and in what follows we denote $\nabla f(s, X_s^{r,x}) = (\nabla f)(s, X_s^{r,x})$). Later on, we occasionally make use of its variant

$$(C.6) \quad \nabla f(s, X_s^{r,x})^T Y(s; r, x) = D_\tau f(s, X_s^{r,x})^T \sigma^{-1}(\tau, X_\tau^{r,x}) Y(\tau; r, x), \quad \forall \tau \in [r, t].$$

Lemma C.1 (Bismut formula). *Suppose f belongs to $C^2(\mathbb{R}^{d+1})$ and suppose f and its derivatives have polynomial growth. Then*

$$(C.7) \quad \mathbb{E}[(\partial_i f)(s, X_s^{r,x})] \\ = \frac{1}{s-r} \mathbb{E} \left[f(s, X_s^{r,x}) \int_r^s [\sigma^{-1}(\tau, X_\tau^{r,x}) Y(\tau; r, x) Y^{-1}(s; r, x)]^{ji} \delta B_\tau^j \right]$$

and

$$(C.8) \quad \partial_i \mathbb{E} f(s, X_s^{r,x}) = \frac{1}{s-r} \mathbb{E} \left[f(s, X_s^{r,x}) \int_r^s [\sigma^{-1}(\tau, X_\tau^{r,x}) Y(\tau; r, x)]^{ji} \delta B_\tau^j \right].$$

Proof. Fix $\tau \in [r, s]$. The identity (C.5) yields

$$\nabla f(s, X_s)^T = [D_\tau f(s, X_s)]^T \sigma^{-1}(\tau) Y(\tau) Y^{-1}(s).$$

Integrating with respect to τ from r to s and taking the expectation give

$$\mathbb{E} \nabla f(s, X_s)^T = \frac{1}{s-r} \mathbb{E} \left[\int_r^s [D_\tau f(s, X_s)]^T \sigma^{-1}(\tau) Y(\tau) Y^{-1}(s) d\tau \right].$$

Formula (C.7) is then followed from the dual relationship (A.6) between the divergence operator δ and the Malliavin derivative D .

To show (C.8), we use (C.6). We integrate with respect to τ from r to s and then take the expectation to obtain

$$\nabla \mathbb{E} f(s, X_s) = \frac{1}{s-r} \mathbb{E} \left[\int_r^s [D_\tau f(s, X_s)]^T \sigma^{-1}(\tau) Y(\tau) d\tau \right].$$

Formula (C.8) follows from the dual relationship (A.6) between δ and D . \square

Lemma C.2. *Suppose that f is differentiable and satisfies*

$$\sup_{s \in [0, T], x \in \mathbb{R}^d} \frac{|f(s, x)|}{1 + |x|^\beta} \leq \kappa$$

for some nonnegative constants κ and β . Then we have

$$(C.9) \quad |\mathbb{E}(\nabla f)(s, X_s^{r,x})| \leq c(T, \Lambda, \lambda) \kappa (1 + |x|^\beta) [1 + (s-r)^{-1/2}]$$

and

$$(C.10) \quad |\nabla \mathbb{E} f(s, X_s^{r,x})| \leq c(T, \Lambda, \lambda) \kappa (1 + |x|^\beta) (s-r)^{-1/2}.$$

Proof. We only provide details for the proof of (C.9). The estimate (C.10) is proved similarly, perhaps in an easier manner. Motivated by the formula (C.7), we first estimate the moment of $\int_r^s [\sigma^{-1}(\tau) Y(\tau) Y^{-1}(s)]^{ji} \delta B_\tau^j$. From (A.8), we see that

$$\begin{aligned} \int_r^s [\sigma^{-1}(\tau) Y(\tau) Y^{-1}(s)]^{ji} \delta B_\tau^j &= \int_r^s [\sigma^{-1}(\tau) Y(\tau)]^{jk} \delta B_\tau^j [Y^{-1}(s)]^{ki} \\ &\quad - \int_r^s [\sigma^{-1}(\tau) Y(\tau)]^{jk} D_\tau^j [Y^{-1}(s)]^{ki} d\tau. \end{aligned}$$

From (C.2) and (C.3), it follows that

$$\sup_{s \in [r, T], x \in \mathbb{R}^d} \mathbb{E} \left| \int_r^s [\sigma^{-1}(\tau) Y(\tau) Y^{-1}(s)]^{ji} \delta B_\tau^j \right|^p \leq c(p, T) [(s-r)^{1/2} + (s-r)]^p.$$

Hence, applying Hölder inequality in (C.7),

$$|\mathbb{E} \nabla f(s, X_s^{r,x})| \lesssim (1 + (s-r)^{-1/2}) [\mathbb{E}(1 + |X_s^{r,x}|^\beta)^2]^{1/2}.$$

Together with (B.4), this completes the proof of (C.9). \square

Proof of Lemma 4.3. Throughout the proof, we denote $\kappa_1 = [\nabla W]_{\beta_1, \infty}$, $\kappa_2 = [\nabla W]_{\beta_2, \alpha}$, $Y = \nabla \varphi$.

Uniqueness: Suppose v is a solution in $C^1([0, T]; C^2(\mathbb{R}^d))$. We apply Itô formula to the process $s \mapsto (v + W)(s, \varphi_s^{r,x})$, taking into account the fact that L_0 is the generator of $\varphi_s^{r,x}$.

$$(C.11) \quad \begin{aligned} d(v + W)(s, \varphi_s^{r,x}) &= (\partial_t + L_0)(v + W)(s, \varphi_s^{r,x}) ds \\ &\quad + \sigma^{ij}(s, \varphi_s^{r,x}) \partial_{x_i}(v + W)(s, \varphi_s^{r,x}) \delta B_s^j. \end{aligned}$$

Since v is a strong solution, we see that $v + W$ satisfies

$$(\partial_t + L_0)(v + W) = L_0 W, \quad (v + W)(T, x) = 0.$$

Thus, integrating (C.11) from r to T yields

$$-(v + W)(r, x) = \int_r^T L_0 W(s, X_s^{r,x}) ds + \int_r^T \sigma^{ij}(s, X_s^{r,x}) \partial_{x_i}(v + W)(s, X_s^{r,x}) \delta B_s^j.$$

Taking expectation in the above identity, we obtain (4.11), which also shows the uniqueness of v .

C^0 -estimate: To prove the estimate (4.12), we write $L_0 W = \partial_i (\frac{1}{2} a^{ij} \partial_j W) + c^j \partial_j W$ where $c^j = -1/2 \partial_i a^{ij}$. Then

$$\mathbb{E} \int_r^T L_0 W(s, X_s^{r,x}) ds = I_1 + I_2$$

where

$$I_1 = \mathbb{E} \int_r^T \partial_i \left(\frac{1}{2} a^{ij}(s, X_s^{r,x}) \partial_j W(s, X_s^{r,x}) \right) ds$$

and

$$I_2 = \mathbb{E} \int_r^T c^j(s, X_s^{r,x}) \partial_j W(s, X_s^{r,x}) ds.$$

It follows from our conditions on L_0 and W that

$$\sup_{t \in [0, T], x \in \mathbb{R}^d} \frac{|a^{ij}(t, x) \partial_i W(t, x)|}{1 + |x|^{\beta_1}} \leq \Lambda \kappa_1 \quad \text{and} \quad \sup_{t \in [0, T], x \in \mathbb{R}^d} \frac{|c^j(t, x) \partial_j W(t, x)|}{1 + |x|^{\beta_1}} \leq \Lambda \kappa_1.$$

Applying Lemma C.2, we obtain

$$I_1 \lesssim \kappa_1 \int_r^T ((s - r)^{-1/2} + 1) ds (1 + |x|^{\beta_1}) \lesssim \kappa_1 [(T - r)^{1/2} + (T - r)] (1 + |x|^{\beta_1}).$$

For the second term, we use (3.7)

$$I_2 \lesssim \kappa_1 \int_r^T \mathbb{E}(1 + |\varphi_s^{r,x}|^{\beta_1}) ds \lesssim \kappa_1 (T - r) (1 + |x|^{\beta_1}).$$

These inequalities altogether imply (4.12).

C^1 -estimate: To show (4.13), we first apply (C.8)

$$\nabla \mathbb{E} L_0 W(s, \varphi_s^{r,x}) = (s - r)^{-1} \mathbb{E}[L_0 W(s, \varphi_s^{r,x}) H(s, x)]$$

where

$$H(s, x) = \int_r^s [\sigma^{-1}(\tau, \varphi_\tau^{r,x}) Y(\tau; r, x)]^T \delta B_\tau.$$

We denote

$$A(\tau, x) = \sigma^{-1}(\tau, X_\tau^{r,x}) Y(\tau; r, x).$$

From (C.5), we see that

$$\partial_{ij}^2 W(s, \varphi_s^{r,x}) = D_\tau^k [\partial_j W(s, \varphi_s^{r,x})] [A(\tau) Y^{-1}(s)]^{ki}, \quad \forall \tau \in [r, s].$$

Thus

$$\begin{aligned} L_0 W(s, \varphi_s^{r,x}) &= \frac{1}{2} a^{ij}(s, X_s^{r,x}) \partial_{ij}^2 W(s, \varphi_s^{r,x}) \\ &= \frac{1}{2} D_\tau^k [\partial_j W(s, \varphi_s^{r,x})] [A(\tau) Y^{-1}(s)]^{ki} a^{ij}(s, X_s^{r,x}) \\ &= \frac{1}{2} (s-r)^{-1} \int_r^s D_\tau^k [\partial_j W(s, \varphi_s^{r,x})] [A(\tau) Y^{-1}(s) a(s, X_s^{r,x})]^{kj} d\tau. \end{aligned}$$

Hence, applying (A.6),

$$\begin{aligned} \partial_l \mathbb{E} L_0 W(s, \varphi_s^{r,x}) &= \frac{1}{2} (s-r)^{-2} \mathbb{E} \int_r^s D_\tau^k [\partial_j W(s, \varphi_s^{r,x})] [A(\tau) Y^{-1}(s) a(s, X_s^{r,x})]^{kj} H^l(s, x) d\tau \\ &= \frac{1}{2} (s-r)^{-2} \mathbb{E} \partial_j W(s, \varphi_s^{r,x}) \int_r^s [A(\tau) Y^{-1}(s) a(s, X_s^{r,x})]^{kj} H^l(s, x) \delta B_\tau^k. \end{aligned}$$

Furthermore, since the random variable

$$G^{jl}(s; r, x) := \int_r^s [A(\tau) Y^{-1}(s) a(s, X_s^{r,x})]^{kj} H^l(s, x) \delta B_\tau^k$$

has mean zero, we can write

$$(C.12) \quad \partial_l \mathbb{E} L_0 W(s, \varphi_s^{r,x}) = \frac{1}{2} (s-r)^{-2} \mathbb{E} [\partial_j W(s, \varphi_s^{r,x}) - \partial_j W(s, x)] G^{jl}(s; r, x).$$

We now estimate the moment $G(s; r, x)$. Applying (A.8), we have

$$\begin{aligned} G^{jl}(s; r, x) &= \int_r^s [A(\tau)]^{km} [Y^{-1}(s) a(s, X_s^{r,x})]^{mj} H^l(s, x) \delta B_\tau^k \\ &= [Y^{-1}(s) a(s, X_s^{r,x})]^{mj} H^l(s, x) \int_r^s [A(\tau)]^{km} \delta B_\tau^k \\ &\quad - \int_r^s D_\tau^k ([Y^{-1}(s) a(s, X_s^{r,x})]^{mj} H^l(s, x)) [A(\tau)]^{km} d\tau. \end{aligned}$$

Using properties of Malliavin derivative, we have

$$\begin{aligned} D_\tau^k ([Y^{-1}(s) a(s, X_s^{r,x})]^{mj} H^l(s, x)) &= D_\tau^k [Y^{-1}(s) a(s, X_s^{r,x})]^{mj} H^l(s, x) + [Y^{-1}(s) a(s, X_s^{r,x})]^{mj} D_\tau^k H^l(s, x). \end{aligned}$$

Hence

$$\begin{aligned} (C.13) \quad G^{jl}(s; r, x) &= [Y^{-1}(s) a(s, X_s^{r,x})]^{mj} H^l(s, x) \int_r^s [A(\tau)]^{km} \delta B_\tau^k \\ &\quad - \int_r^s D_\tau^k [Y^{-1}(s) a(s, X_s^{r,x})]^{mj} H^l(s, x) [A(\tau)]^{km} d\tau \\ &\quad - \int_r^s [Y^{-1}(s) a(s, X_s^{r,x})]^{mj} D_\tau^k H^l(s, x) [A(\tau)]^{km} d\tau. \end{aligned}$$

Since a belongs to C_b^2 , estimate (C.3) is valid, the moments of $A(\tau)$ is also uniformly bounded (because a is strictly elliptic), and all the terms appear in G^{jl} has finite

moments of all orders. In addition, observe that

$$D_\tau^i H^l(s, x) = 1_{\{r \leq \tau\}} A(\tau)^{il} + \int_r^s D_\tau^i A(u)^{kl} \delta B_u^k.$$

Thus, the L^p -norm of $H(s, x)$ and $DH(s, x)$ will contribute a factor $(r - s)^{1/2}$. Therefore, it follows from Burkholder-Davis-Gundy inequality and Hölder inequality that

$$(C.14) \quad \sup_{x \in \mathbb{R}^d} \|G^{jl}(s; r, x)\|_p \leq c(p, \lambda, \Lambda)[(s - r) + (s - r)^{3/2}], \forall p \geq 1.$$

Using the Hölder continuity of W , for every $p \geq 1$, we have

$$\|\nabla W(s, \varphi_s^{r,x}) - \nabla W(s, x)\|_p \leq \kappa_2(1 + |\varphi_s^{r,x}|^{\beta_2} + |x|^{\beta_2})|\varphi_s^{r,x} - x|^\alpha \|p.$$

Taking into account the moment estimate (B.4) and Hölder inequality, this gives

$$(C.15) \quad \|\nabla W(s, \varphi_s^{r,x}) - \nabla W(s, x)\|_p \leq c(\alpha, \beta_2, p, \Lambda) \kappa_2(1 + |x|^{\beta_2})(s - r)^{\alpha/2}.$$

Thus, applying Cauchy-Schwartz inequality in (C.12) yields

$$|\partial_l \mathbb{E} L_0 W(s, \varphi_s^{r,x})| \leq c(\lambda, \Lambda)(s - r)^{-2} \|\nabla W(s, \varphi_s^{r,x}) - \nabla W(s, x)\|_2 \|G(s; r, x)\|_2.$$

Applying the moment estimate for G and (C.15), we obtain

$$|\partial_l \mathbb{E} L_0 W(s, \varphi_s^{r,x})| \leq c(\lambda, \Lambda)[(s - r)^{\alpha/2-1} + (s - r)^{\alpha/2-1/2}] \kappa_2(1 + |x|^{\beta_2}),$$

which together with (4.11) implies (4.13)

$C^{1,\alpha'}$ -estimate: This is the only place where we use the fact that the second derivatives of a are Hölder continuous. Each term appeared on the right hand side (C.13) is either differentiable or Hölder continuous in the x -variable. Thus, we obtain easily the estimate

$$(C.16) \quad \|G(s; r, x) - G(s; r, y)\|_p \leq c(p, \lambda, \Lambda)[(s - r) + (s - r)^{3/2}]|x - y|^\alpha.$$

From (C.15), we see that

$$\begin{aligned} & \|\nabla W(s, \varphi_s^{r,x}) - \nabla W(s, x) - \nabla W(s, \varphi_s^{r,y}) + \nabla W(s, y)\|_p \\ & \leq \|\nabla W(s, \varphi_s^{r,x}) - \nabla W(s, x)\|_p + \|\nabla W(s, \varphi_s^{r,y}) - \nabla W(s, y)\|_p \\ & \leq c(\alpha, p, \Lambda) \kappa_2(1 + |x|^{\beta_2} + |y|^{\beta_2})(s - r)^{\alpha/2}. \end{aligned}$$

On the other hand, we also have

$$\begin{aligned} & \|\nabla W(s, \varphi_s^{r,x}) - \nabla W(s, x) - \nabla W(s, \varphi_s^{r,y}) + \nabla W(s, y)\|_p \\ & \leq \|\nabla W(s, \varphi_s^{r,x}) - \nabla W(s, \varphi_s^{r,y})\|_p + \|\nabla W(s, x) - \nabla W(s, y)\|_p \\ & \leq \kappa_2(W) \|(1 + |\varphi_s^{r,x}|^{\beta_2} + |\varphi_s^{r,y}|^{\beta_2})|\varphi_s^{r,x} - \varphi_s^{r,y}|^\alpha\|_p \\ & \quad + \kappa_2(W)(1 + |x|^{\beta_2} + |y|^{\beta_2})|x - y|^\alpha \\ & \leq c(\alpha, p, \Lambda) \kappa_2(1 + |x|^{\beta_2} + |y|^{\beta_2})|x - y|^\alpha, \end{aligned}$$

where the last estimate comes from (B.4) and that fact that the derivative of the map $x \mapsto \varphi_s^{r,x}$ has finite moments uniformly in x . Interpolating these two inequalities we obtain

$$(C.17) \quad \begin{aligned} & \|\nabla W(s, \varphi_s^{r,x}) - \nabla W(s, x) - \nabla W(s, \varphi_s^{r,y}) + \nabla W(s, y)\|_p \\ & \leq c(\alpha, p, \Lambda) \kappa_2(1 + |x|^{\beta_2} + |y|^{\beta_2})|x - y|^{\vartheta\alpha} (s - r)^{(1-\vartheta)\alpha/2} \end{aligned}$$

for any $\vartheta \in [0, 1]$. Thus, from (C.12), applying Cauchy-Schwartz inequality we see that

$$\begin{aligned} & |\nabla \mathbb{E} L_0 W(s, \varphi_s^{r,x}) - \nabla \mathbb{E} L_0 W(s, \varphi_s^{r,y})| \\ & \leq (s-r)^{-2} \|\nabla W(s, \varphi_s^{r,x}) - \nabla W(s, x) - \nabla W(s, \varphi_s^{r,y}) + \nabla W(s, y)\|_2 \|G(s; r, x)\|_2 \\ & \quad + (s-r)^{-2} \|\nabla W(s, \varphi_s^{r,y}) - \nabla W(s, y)\|_2 \|G(s; r, x) - G(s; r, y)\|_2. \end{aligned}$$

Using (C.17), (C.15), (C.14) and (C.16), we obtain

$$\begin{aligned} & |\nabla \mathbb{E} L_0 W(s, \varphi_s^{r,x}) - \nabla \mathbb{E} L_0 W(s, \varphi_s^{r,y})| \\ & \leq c(\alpha, \lambda, \Lambda) \kappa_2 (1 + |x|^{\beta_2} + |y|^{\beta_2}) |x - y|^{\vartheta \alpha} [(s-r)^{(1-\vartheta)\alpha/2-1} + (s-r)^{(1-\vartheta)\alpha/2-1/2}] \\ & \quad + c(\alpha, \lambda, \Lambda) \kappa_2 (1 + |y|^{\beta_2}) |x - y|^\alpha [(s-r)^{\alpha/2-1} + (s-r)^{\alpha/2-1/2}]. \end{aligned}$$

Therefore, choosing $\vartheta < 1$, this estimate together with (4.11) implies (4.14). \square

Acknowledgment. The authors sincerely thank the referee for the many constructive and detailed comments which were of great help in revising the manuscript. We also thank Samy Tindel and Massimiliano Gubinelli for their interest in our paper.

REFERENCES

- [1] E. Alòs and D. Nualart. Stochastic integration with respect to the fractional Brownian motion. *Stoch. Stoch. Rep.*, 75(3):129–152, 2003.
- [2] L. Ambrosio. Transport equation and Cauchy problem for BV vector fields. *Invent. Math.*, 158(2):227–260, 2004.
- [3] H. Bahouri, J.-Y. Chemin, and R. Danchin. *Fourier analysis and nonlinear partial differential equations*, volume 343 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, 2011.
- [4] G. Ben Arous and M. Ledoux. Grandes déviations de freidlin-wentzell en norme hölderienne. *Séminaire de Probabilités XXVIII, Lecture Notes in Math., 1583 (1994), Springer*, pages 293–299.
- [5] R. Catellier and M. Gubinelli. Averaging along irregular curves and regularisation of odes. *arXiv preprint arXiv:1205.1735*, 2012.
- [6] S. Cerrai. *Second order PDE's in finite and infinite dimension*, volume 1762 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2001. A probabilistic approach.
- [7] K. Chouk and M. Gubinelli. Nonlinear pdes with modulated dispersion. *arXiv preprint arXiv:1303.0822*, 2013.
- [8] K. Chouk and M. Gubinelli. Nonlinear pdes with modulated dispersion ii: Korteweg–de vries equation. *arXiv preprint arXiv:1406.7675*, 2014.
- [9] G. Crippa and C. De Lellis. Estimates and regularity results for the DiPerna-Lions flow. *J. Reine Angew. Math.*, 616:15–46, 2008.
- [10] A. M. Davie. Uniqueness of solutions of stochastic differential equations. *Int. Math. Res. Not. IMRN*, (24):Art. ID rnm124, 26, 2007.
- [11] P. Del Moral. *Feynman-Kac formulae*. Probability and its Applications (New York). Springer-Verlag, New York, 2004. Genealogical and interacting particle systems with applications.
- [12] P. Del Moral. *Mean field simulation for Monte Carlo integration*, volume 126 of *Monographs on Statistics and Applied Probability*. CRC Press, Boca Raton, FL, 2013.
- [13] R. J. DiPerna and P.-L. Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.*, 98(3):511–547, 1989.
- [14] B. K. Driver and Y. Hu. On heat kernel logarithmic Sobolev inequalities. In *Stochastic analysis and applications (Powys, 1995)*, pages 189–200. World Sci. Publ., River Edge, NJ, 1996.
- [15] R. M. Dudley. Sample functions of the Gaussian process. *Ann. Probability*, 1(1):66–103, 1973.
- [16] X. Fernique. Régularité des trajectoires des fonctions aléatoires gaussiennes. In *École d'Été de Probabilités de Saint-Flour, IV-1974*, pages 1–96. Lecture Notes in Math., Vol. 480. Springer, Berlin, 1975.

- [17] D. Feyel and A. de La Pradelle. Curvilinear integrals along enriched paths. *Electron. J. Probab.*, 11:no. 34, 860–892 (electronic), 2006.
- [18] F. Flandoli, M. Gubinelli, and E. Priola. Flow of diffeomorphisms for SDEs with unbounded Hölder continuous drift. *Bull. Sci. Math.*, 134(4):405–422, 2010.
- [19] F. Flandoli, M. Gubinelli, and E. Priola. Well-posedness of the transport equation by stochastic perturbation. *Invent. Math.*, 180(1):1–53, 2010.
- [20] M. Freidlin. *Functional integration and partial differential equations*. Annals of Mathematics Studies, 109. Princeton University Press, Princeton, NJ, 1985.
- [21] A. M. Garsia, E. Rodemich, and H. Rumsey, Jr. A real variable lemma and the continuity of paths of some Gaussian processes. *Indiana Univ. Math. J.*, 20:565–578, 1970/1971.
- [22] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.
- [23] M. Grigoriu. *Stochastic systems. Uncertainty quantification and propagation*. Springer Series in Reliability Engineering. Springer-Verlag, London, 2012.
- [24] M. Gubinelli. Controlling rough paths. *J. Funct. Anal.*, 216(1):86–140, 2004.
- [25] M. Gubinelli and S. Tindel. Rough evolution equations. *Ann. Probab.*, 38(1):1–75, 2010.
- [26] T. Hida. *Brownian motion*, volume 11 of *Applications of Mathematics*. Springer-Verlag, New York-Berlin, 1980. Translated from the Japanese by the author and T. P. Speed.
- [27] H. Holden and Y. Hu. Finite difference approximation of the pressure equation for fluid flow in a stochastic medium—a probabilistic approach. *Comm. Partial Differential Equations*, 21(9-10):1367–1388, 1996.
- [28] Y. Hu. Itô-Wiener chaos expansion with exact residual and correlation, variance inequalities. *J. Theoret. Probab.*, 10(4):835–848, 1997.
- [29] Y. Hu, J. Huang, D. Nualart, and S. Tindel. Stochastic heat equations with general multiplicative gaussian noises: Hölder continuity and intermittency. *arXiv preprint arXiv:1402.2618*, 2014.
- [30] Y. Hu and K. Le. A multiparameter Garsia-Rodemich-Rumsey inequality and some applications. *Stochastic Process. Appl.*, 123(9):3359–3377, 2013.
- [31] Y. Hu and K. Le. Nonlinear young integrals via fractional calculus. *arXiv preprint arXiv:1503.00328*, 2015.
- [32] Y. Hu, F. Lu, and D. Nualart. Feynman-Kac formula for the heat equation driven by fractional noise with Hurst parameter $H < 1/2$. *Ann. Probab.*, 40(3):1041–1068, 2012.
- [33] Y. Hu and D. Nualart. Differential equations driven by Hölder continuous functions of order greater than $1/2$. In *Stochastic analysis and applications*, volume 2 of *Abel Symp.*, pages 399–413. Springer, Berlin, 2007.
- [34] Y. Hu and D. Nualart. Rough path analysis via fractional calculus. *Trans. Amer. Math. Soc.*, 361(5):2689–2718, 2009.
- [35] Y. Hu, D. Nualart, and J. Song. Feynman-Kac formula for heat equation driven by fractional white noise. *Ann. Probab.*, 39(1):291–326, 2011.
- [36] N. V. Krylov and E. Priola. Elliptic and parabolic second-order PDEs with growing coefficients. *Comm. Partial Differential Equations*, 35(1):1–22, 2010.
- [37] H. Kunita. *Stochastic flows and stochastic differential equations*, volume 24 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1990.
- [38] M. Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.
- [39] T. Lyons. Differential equations driven by rough signals. I. An extension of an inequality of L. C. Young. *Math. Res. Lett.*, 1(4):451–464, 1994.
- [40] T. J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, 14(2):215–310, 1998.
- [41] M. B. Marcus and J. Rosen. *Markov processes, Gaussian processes, and local times*, volume 100 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.
- [42] M. M. Meerschaert, W. Wang, and Y. Xiao. Fernique-type inequalities and moduli of continuity for anisotropic Gaussian random fields. *Trans. Amer. Math. Soc.*, 365(2):1081–1107, 2013.
- [43] D. Nualart. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.

- [44] C. Preston. Banach spaces arising from some integral inequalities. *Indiana Univ. Math. J.*, 20:997–1015, 1970/1971.
- [45] C. Preston. Continuity properties of some Gaussian processes. *Ann. Math. Statist.*, 43:285–292, 1972.
- [46] B. Simon. *Functional integration and quantum physics*. Second edition. AMS Chelsea Publishing, Providence, RI, 2005.
- [47] A. V. Skorohod. On a generalization of the stochastic integral. *Theory of Probability and its Applications*, 20:219, 1976.
- [48] D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition.
- [49] M. Talagrand. *The generic chaining*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2005.
- [50] J. Xin. *An introduction to fronts in random media*. Surveys and Tutorials in the Applied Mathematical Sciences, 5. Springer-Verlag, New York, 2009.
- [51] L. C. Young. An inequality of the Hölder type, connected with Stieltjes integration. *Acta Math.*, 67(1):251–282, 1936.

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF KANSAS, LAWRENCE, KANSAS, 66045
E-mail address: yhu@ku.edu, khoale@ku.edu